Università degli Studi di Napoli L'Orientale

La comunicazione parlata 3

Atti del terzo congresso internazionale
del Gruppo di Studio sulla Comunicazione Parlata

volume II

A cura di
Massimo Pettorino
Antonella Giannini
Francesca M. Dovetto

Società di Linguistica Italiana
Gruppo di Studio sulla Comunicazione Parlata

La comunicazione parlata 3

Atti del congresso internazionale
(Napoli, 23-25 febbraio 2009)

Volume II

a cura di M. Pettorino, A. Giannini, F.M. Dovetto

Università degli Studi di Napoli L'Orientale

Università degli Studi di Napoli L'Orientale

Questa opera è protetta dalla Legge sul diritto d'autore. Tutti i diritti sono riservati, anche nel caso di utilizzo parziale.
La riproduzione di questa opera, anche se parziale o in copia digitale, è ammessa solo ed esclusivamente nei limiti stabiliti dalla Legge.

© 2010 by Università degli Studi di Napoli L'Orientale, Napoli

ISBN: 978-88-95044-83-5

Copertina: progetto grafico di Erica Carol Tortora

Indice del Volume II

Parte III - *La dimensione fonica*

3 ***L'espansione della dittongazione nei giovani pescatori di Pozzuoli (NA). Dati acustici su un cambiamento fonetico in corso***
Giovanni Abete, Adrian Simpson

23 ***Analisi prosodica della cronaca in diretto del giornalista sportivo***
Georges Boulakia, Catherine Mathon, Takeki Kamiyama

45 ***Griglie tonali***
Amedeo De Dominicis

71 ***Il parlato dei mass media: analisi multilingue del parlato dei telegiornali***
Antonella Giannini, Massimo Pettorino

85 ***Il ruolo dell'intonazione nell'interpretazione dell'enunciato: informazioni categoriali o categorie informative?***
Barbara Gili Fivela

109 ***Lingua parlata – lingua cantata***
Cantabilità delle lingue: rivisitazione di un topos – analisi storico-culturale, dimensione linguistico-tipologica e analisi linguistico-musicologica
Elmar Schafroth

127 ***Sistemi fonetici in contatto***
La variabilità di /r/ nell'italiano di tedescofoni altoatesini
Lorenzo Spreafico, Alessandro Vietti

149 ***The pragmatic ground of flores rhetoricae***
Carlotta Viti

Parte IV - *La dimensione morfosintattica*

169 ***I focalizzatori in un corpus di parlato spontaneo (C-ORAL-ROM Italia)***
Daniela Acciardi

191 ***La dislocazione a destra tra italiano comune e variazione regionale***
Claudia Crocco

211 ***Predicazione nominale negli usi dell'italiano parlato***
L'esempio della classe delle forme lessicali di "sentimento"
Emilio D'Agostino

227 ***Costrutti con c'è e nome proprio in una telecronaca sportiva: configurazioni funzionali e valori testuali***
Nunzio La Fauci, Heike Necker, Sophia Simon, Liana Tronci

241 ***Le espressioni di sentimento nel LIP***
Alberto Maria Langella, Simona Messina, Serena Santonicola

259 ***L'immediatezza comunicativa nelle lingue romanze***
Figure di sintassi nel C-ORAL-ROM
Araceli López Serena

281 ***Spontaneous speech intonation in Italian***
An experimental analysis with a macrosyntactic view
Philippe Martin

Parte V - *Antropologia e sociologia della comunicazione parlata*

295 ***Le formulazioni in contesti istituzionali: un'analisi linguistica e sociologica***
Claudio Baraldi

315 ***Il parlato nella storia dell'antropologia statunitense. Un progetto politico***
Alessandro Duranti

335 ***La lingua parlata del teatro di narrazione***
Analisi di tre monologhi (Paolini, Celestini, Enia)
Carmela Lucia

361 ***L'improvvisazione in ottava rima in Toscana***
Una pratica di comunicazione "solenne"
Grazia Tiezzi

Parte VI - Comunicazione parlata e patologie del linguaggio

385 ***Ipotesi per una valutazione dei deficit afasici all'interfaccia tra sintassi e scelte pragmatico-discorsive***
Camilla Bettoni, Maria Elena Favilla, Lucia Ferroni

403 ***Il parlato schizofrenico***
Una proposta di analisi conversazionale per seguire le piste topicali
Annamaria Cacchione

425 ***La produzione verbale dei soggetti con Sindrome di Down: aspetti fonologici e morfologici***
Patrizia Sorianello

Parte VII - Strumenti e tecniche per lo studio della comunicazione parlata

457 ***Il processo di Erba tra Analisi Transazionale e Counselling giuridico***
Riflessioni per una comunicazione efficace
Monica Dell'Aglio, Maria Lasaponara

477 ***Towards Automatic Dialogue Understanding***
Rodolfo Delmonte, Antonella Bristot, Vincenzo Pallotta

507 ***Normalizing speech transcriptions for Natural Language Processing***
Anne Dister, Matthieu Constant, Gérald Purnelle

521 ***È possibile "guardare la comunicazione?"***
Esperienze di dialogo tra Analisi della Conversazione e video osservazione dei processi comunicativi
Federico Farini

541 ***Sbagliando si impara***
Strategie per l'attenuazione dell'errore in un sistema di riconoscimento speaker independent per la lingua italiana
Maria Palmerini

565 ***I gesti interazionali come risorsa per la gestione dell'alternanza dei turni***
Maria Eleonora Sciubba

589 ***La costruzione di una rappresentazione condivisa attraverso l'impiego del focus group***
Patrizia Velotti, Cecilia Serena Pace, Giulio Cesare Zavattini

603 ***Parlare italiano: towards a multidimensional description and a multidisciplinary explanation***
Miriam Voghera

PARTE III

LA DIMENSIONE FONICA

L'espansione della dittongazione nei giovani pescatori di Pozzuoli (NA).

Dati acustici su un cambiamento fonetico in corso

Giovanni Abete, Adrian Simpson
Friedrich-Schiller-Universität Jena
giovanni.abete@libero.it; adrian.simpson@uni-jena.de

Abstract

The dialect of Pozzuoli is characterized by a context-dependent alternation between monophthongs and diphthongs. A word such as pisci *'fishes' can be rendered as [pʏɪʃʃ] before high prosodic boundaries, while it will be rendered as [piʃʃ] in internal position. This phenomenon is very pervasive, because it involves 4 vocalic variables ((i), (e), (o), (u)) and is unlimited by syllable structure. In this paper we will discuss how the younger fishermen have been spreading this phonetic alternation to the variable (ɛ). The data were collected from casual conversations with fishermen of Pozzuoli. An acoustic analysis was performed. We measured the duration of the vocalic segments and used several techniques to characterize the diphthong dynamics. The diffusion of diphthongization in this community is particularly intriguing because of the sociolinguistic status of diphthongs. They are overtly stigmatized, even by those speakers who actually extend the use of diphthongization themselves. The results of this research, therefore, highlight the question of the complex interactions between opinions, attitudes and actual linguistic behaviour. These data suggest the search for "covert" factors in order to explain the maintenance and spreading of non-standard locally-marked features.*

1 Introduzione[1]

Si presentano in questo contributo alcuni risultati di uno studio sulla "dittongazione spontanea" nei dialetti dell'Italia meridionale[2]. Con "dittongazione spontanea" si intende qui un fenomeno di alternanza sincronica tra esiti monottongali e esiti dittongali di alcune variabili vocaliche. Le analisi acustiche hanno mostrato come tale alternanza

[1] Ai fini accademici i paragrafi 1-4 sono da attribuirsi a Adrian Simpson e 5-8 a Giovanni Abete.
[2] Cfr. Abete (in preparazione).

dipenda strettamente dalla posizione della variabile nella struttura prosodica: le varianti dittongali sono in genere limitate alla posizione finale di costituenti prosodici alti, come il sintagma intonativo, mentre le varianti monottongali compaiono nelle posizioni interne[3]. Rispetto a questa tendenza generale, ciascuno dei dialetti esaminati presenta delle peculiarità, sia in riferimento al numero e al tipo delle variabili vocaliche coinvolte, sia relativamente ai *patterns* di variazione, sia in rapporto al diverso significato sociolinguistico di volta in volta assunto dalle varianti dittongali (cfr. Abete, 2008).

In questa sede ci si concentrerà su alcuni dati relativi al dialetto di Pozzuoli, che risultano di particolare interesse per la questione della diffusione di tale fenomeno attraverso le generazioni. Come si cercherà di dimostrare, nella comunità dei pescatori di Pozzuoli la dittongazione è limitata nei parlanti anziani alle vocali toniche alte e medio-alte, mentre nei giovani il fenomeno si è esteso anche alla vocale anteriore medio-bassa. I dati che verranno presentati sono il frutto di analisi acustiche e di metodologie quantitative, applicate a parlato dialettale elicitato in situazioni conversazionali piuttosto naturali[4].

Il fenomeno è di particolare interesse anche in rapporto al peculiare status sociolinguistico delle varianti dittongali. L'uso di tali varianti, infatti, è fortemente stigmatizzato, sia da parte degli abitanti dei comuni limitrofi, sia da parte degli stessi puteolani, che le ritengono una marca di bassa estrazione sociale[5]. Anche i parlanti che presentano il fenomeno in maniera consistente, come i pescatori oggetto di questa analisi, esprimono giudizi fortemente negativi nei confronti di tali tratti e, in condizioni di maggiore formalità, tendono a prediligere l'uso di varianti monottongali. Eppure, negli usi privati e più spontanei, la dittongazione continua a resistere e, come si vedrà, conquista addirittura nuove porzioni del lessico.

Oggetto di questo studio è quindi l'espansione dia-generazionale di un fenomeno fonetico fortemente e apertamente stigmatizzato dagli stessi parlanti che ne estendono l'uso. In § 2 verrà introdotto il fenomeno in

[3] Cfr. Abete and Simpson (in press). Per il dialetto di Pozzuoli una stretta relazione tra dittongazione e posizione nella frase era stata già rilevata da Rohlfs (1966: § 12).

[4] Le registrazioni sono state effettuate con una versione modificata dell'intervista libera (cfr. Como, 2006). Per maggiori dettagli sulla procedura di raccolta si veda Abete and Simpson (in press: § 3).

[5] Cfr. Abete (2008).

esame, in § 3 verrà descritta la parte del corpus relativa al presente contributo, in § 4 si forniranno le specifiche tecniche per l'analisi acustica dei dittonghi, in § 5 verrà delineato un quadro d'insieme delle varianti nello spazio vocalico, in § 6 si presenteranno i dati relativi alla variabile (ɛ) e si metterà in evidenza l'espansione della dittongazione nei parlanti più giovani, in § 7 si descriveranno alcune particolarità nella diffusione lessicale del fenomeno, in § 8 si discuteranno brevemente i risultati della ricerca.

2 L'alternanza monottongo/dittongo

Il dialetto di Pozzuoli è caratterizzato da un fenomeno di alternanza sincronica tra esiti monottongali ed esiti dittongali di determinate variabili vocaliche. Tale alternanza è condizionata dalla posizione della variabile nella struttura prosodica e le varianti dittongali sono limitate in genere alla posizione finale di sintagma intonativo. Di seguito si forniscono alcuni esempi del fenomeno in questione:

A1.	ɪ p̰**i**ʃː i na ˈvɔt̰ə ‖ 'i pesci di una volta'	B1.	ˈpɪɟːə ˈβɔɡʊ b̰ɤʃː ‖(.) 'pigli pochi pesci'
A2.	**ib̰i**ʃː ɪ na ɔdə 'i pesci di una volta'	B2.	e ˈzɛmə b̰iˈʎːad̰ iː b̰**ɤ̰ɪ̰**ʃʃ ‖(.) 'hai sempre pigliato i pesci'
A3.	tɾə̣ pːɪʃkaˈd̰**ʊ**ɾə βə̰ pːʌːht^{h} ‖(.) 'tre pescatori per parte'	B3.	i pːɪʃkəˈt̰ɪ**ʊ** ‖ zoŋɡ ɐˈsːɑɡ̰ ‖(.) 'i pescatori sono assai'
A4.	ɪ p̰ɪʃkaˈd**ʊ**ɾ ɪ lːɡ̰ ‖(.) 'i pescatori di là'	B4.	i b̰**i**ʃkɐˈdɪʊɾ ‖ iː ɾə̰tːsaˈjʊːlə̣ ‖ 'i pescatori... i *rezzaioli*...'

Nella colonna di sinistra le parole *pesci* e *pescatori* sono realizzate in posizione interna e presentano esiti monottongali delle vocali toniche del tipo [i] e [ʊ], mentre nella colonna di destra queste stesse parole sono realizzate in posizione finale di sintagma intonativo e presentano esiti dittongali del tipo [ɤɪ] e [ɪʊ]. Il fenomeno coinvolge nel dialetto degli anziani le vocali alte e medio-alte anteriori e posteriori e non conosce limitazioni relativamente alla natura aperta o chiusa della sillaba e alla struttura accentuale della parola[6].

[6] Per una descrizione più dettagliata e per ulteriori esempi si rinvia ad Abete (2006).

Come si mostrerà in questo contributo, questo fenomeno si è esteso in un gruppo di giovani pescatori, di età inferiore ai quarant'anni, alla vocale anteriore medio-bassa, instaurando *patterns* di variazione del tutto analoghi a quelli già in vigore per le vocali alte e medio-alte. Si vedano ad esempio le realizzazioni della parola *finendo* in questa sequenza realizzata da un pescatore di 28 anni: [stʌ fuˈnːɛnːə̯ qːɑə ‖ stʌ ʃʊˈnːɛ̃nːʌ lːʸɑː ‖] 'sta finendo qua, sta finendo là...'; in opposizione a [ʊ ˈmwaɾə sta ʃʊˈnːaẹ̯nːə ‖] 'il mare sta finendo', realizzata dallo stesso parlante. In tal caso il fenomeno di alternanza riguarda la variabile (ɛ), con esiti di tipo monttongale [ɛ] o [ɛ̃] in posizione interna, e esiti di tipo dittongale [aẹ̯] in posizione finale.

3 I parlanti e il corpus

Il materiale utilizzato per questa ricerca costituisce un sottoinsieme del corpus presentato in Abete (in preparazione) e consta di 6 ore di registrazioni di parlato spontaneo di 8 parlanti maschi, 4 dei quali hanno età comprese tra i 26 e i 36 anni, mentre i restanti 4 hanno età comprese tra i 48 e i 67 anni. Gli informatori sono tutti pescatori, hanno istruzione piuttosto bassa, e provengono dai quartieri Monteruscello e Rione Toiano, ossia dai luoghi che hanno accolto la maggioranza degli sfollati del centro storico di Pozzuoli, dopo il suo sgombero a causa del bradisismo[7]. Questi parlanti costituiscono un gruppo sociale molto compatto: alcuni di essi sono legati da vincoli di parentela, altri lavorano insieme sulle stesse barche, tutti condividono una parte considerevole della giornata in una piccola area di rimessaggio, dove si intrattengono a rammendare le reti e a chiacchierare. Le caratteristiche principali del campione sono riportate in tabella 1[8].

Come evidenziato, i parlanti possono essere divisi in due gruppi: i giovani, di età inferiore ai 40 anni, e gli anziani, di età superiore ai 40 anni. Questa distinzione non era stata imposta preliminarmente alla scelta del campione, ma è emersa invece dall'analisi dei materiali e in particolare dal comportamento nettamente distinto che i due gruppi mostrano relativamente alla variabile (ɛ). "Giovani" e "anziani" sono dunque etichette di comodo, che servono a identificare questi due gruppi, e non si riferiscono a categorie sociologicamente fondate.

[7] Cfr. Giglia (1997).

[8] I nomi reali sono stati sostituiti con nomi di fantasia.

	NOME	ETÀ	PROFESSIONE	ISTRUZIONE
giovani	Raffaele	26	pescatore	III media
	Giuseppe	33	pescatore	III media
	Gennaro	34	pescatore	III media
	Gabriele	36	pescatore	III media
anziani	Ciro	48	pescatore	III media
	Salvatore	50	pescatore	IV elementare
	Alberto	61	pescatore	V elementare
	Antonio	67	pescatore	analfabeta

Tabella 1 Gruppo dei giovani e gruppo degli anziani nel corpus di Pozzuoli

Il corpus è stato segmentato ed etichettato in maniera manuale, limitatamente a una lista di parole precedentemente selezionate. La procedura è partita da un ascolto impressionistico dei materiali raccolti, quindi dalla scelta di determinati *items* lessicali che presentavano variabilità degli esiti nelle vocali toniche, in particolare variabilità tra realizzazioni di tipo dittongale e realizzazioni di tipo monottongale. Sulla base degli *items* scelti sono state individuate tutte le loro realizzazioni nel corpus. L'etichettatura ha previsto diversi livelli: segmenti, parole, sintagmi intonativi, enunciati. Altre informazioni prosodiche sono state inserite nella trascrizione al livello segmentale. Tale procedura ha consentito di effettuare un'analisi delle variabili fonetiche all'interno di determinati tipi lessicali, controllando la posizione che le variabili occupano in strutture prosodiche gerarchicamente più alte, come il sintagma intonativo, e fornendo informazioni sul contesto discorsivo più ampio, grazie alla trascrizione del parlato contenuto in uno o più enunciati.

Come accennato, l'alternanza monottongo/dittongo è un fenomeno che dipende in primo luogo dalla posizione della variabile vocalica nella struttura prosodica. Nella presente analisi è risultata di particolare importanza l'opposizione tra la posizione interna al sintagma intonativo[9] e la posizione finale di sintagma intonativo, comunemente definita "prepausale". All'interno della posizione finale è parso opportuno operare una ulteriore suddivisione tra enunciati che terminano con un'intonazione ascendente (esprimendo ad esempio l'idea di "continuazione") e enunciati privi di tale contorno

[9] Il sintagma intonativo viene qui definito come la sequenza di parlato inclusa tra due pause prosodiche di livello superiore (cfr. Nespor, 1993: 205 ss.)

intonativo[10]. La posizione finale non marcata è stata dunque etichettata come *Finale A*, la posizione finale con intonazione ascendente è stata etichettata come *Finale B*. I dati del lavoro di etichettatura sono stati quindi trasferiti in un data-base attraverso una serie di procedure automatizzate.
Per la parte relativa al campione Pozzuoli, l'etichettatura ha individuato un totale di 944 *tokens*, che sono stati sottoposti a diversi tipi di analisi acustiche. Di questi *tokens*, 344 sono realizzazioni della variabile vocalica tonica anteriore medio-bassa (ɛ). È specificatamente su questo materiale che si incentrerà il presente contributo.

GIOVANI		ANZIANI	
parlante	tokens di (ɛ)	parlante	tokens di (ɛ)
Ciro	42	Raffaele	39
Alessandro	27	Francesco	51
Mario	65	Biagio	34
Salvatore	46	Luigi	30
subtot.	**180**	subtot.	**154**

Tabella 2: Numero di *tokens* di (ɛ) per parlanti e per gruppi (tot. = 344)

TIPI LESS.	N. TOKENS	GIOVANI	ANZIANI
metro	46	35	11
-mente	46	26	20
-ello	42	29	13
sempre	40	12	28
terra	39	21	18
-etta	22	11	11
bello	21	11	10
toponimi	19	4	15

Tabella 3: Tipi lessicali più frequenti nel gruppo dei giovani e in quello degli anziani

Nella tabella 2 si riporta il numero dei *tokens* di (ɛ) analizzati per ciascun parlante del gruppo dei giovani e del gruppo degli anziani. Nella tabella 3 vengono presentati gli *items* lessicali più frequenti in questa parte del corpus. Alcuni *items* sono stati raggruppati in una stessa categoria, in base alla presenza di un particolare suffisso (es. -

[10] Analisi statistiche hanno evidenziato l'importanza in più dialetti di questa distinzione, sia relativamente ai *patterns* di allungamento, sia relativamente all'ampiezza dei movimenti dittongali (cfr. Abete and Simpson, in press: § 5.2).

ello, *-mente*, *-etto*). Come si mostrerà in seguito (§ 7), una categoria speciale è costituita dai toponimi[11], che sembrano essere interessati dalla dittongazione in una maniera peculiare.

4 Caratterizzazione dei dittonghi

L'analisi dell'alternanza tra realizzazioni monottongali e realizzazioni dittongali è stata condotta su base acustica, sviluppando un metodo per la caratterizzazione delle traiettorie dittongali e un indice numerico dell'ampiezza di tali traiettorie. Il metodo di caratterizzazione della traiettoria dittongale costituisce una evoluzione del metodo di Holbrook and Fairbancks (1962) e segue essenzialmente Simpson (1998). Le traiettorie formantiche vengono descritte da un insieme di misurazioni di numero variabile (in rapporto alla lunghezza del segmento vocalico), prese secondo un intervallo di tempo costante. Di ciascun segmento vocalico vengono scartati in automatico i primi 20 ms. e gli ultimi 20 ms.; ciò è necessario per ridurre l'impatto delle transizioni formantiche, che nei criteri di segmentazione utilizzati in questa ricerca vengono incluse nel segmento vocalico. La prima e l'ultima misurazione delle formanti vengono effettuate subito dopo i primi 20 ms. e subito prima degli ultimi 20 ms. I restanti punti in cui effettuare le misurazioni vengono calcolati dividendo in parti uguali la porzione tra la prima e l'ultima misurazione, facendo in modo che la lunghezza dell'unità di segmentazione sia la più vicina possibile al valore nominale di 20 ms. Quindi vengono stimati i valori delle prime tre formanti per ciascuno dei punti precedentemente individuati. Ogni segmento vocalico inferiore ai 60 ms. viene trattato tecnicamente come un monottongo e viene caratterizzato da un singolo insieme di valori di F1, F2 e F3 presi nella porzione centrale della vocale, al centro di una finestra di 20 ms.

Oltre a consentire varie rappresentazioni grafiche delle traiettorie dittongali (cfr. Abete and Simpson, in press), la procedura descritta fornisce i dati per il calcolo di un indice numerico dell'ampiezza dei movimenti dittongali. L'indice misura la distanza euclidea nello spazio F1-F2 che intercorre tra i due timbri più distanti raggiunti dal

[11] I toponimi che presentano una realizzazione di (ɛ) tonica in questa parte del corpus sono *Civitavecchia*, *Maddalena*, *Marinella*, *Port'Ercole, Ventotene*.

segmento nella transizione dittongale. La base della procedura di calcolo sta nell'individuazione degli scarti tra valore massimo e valore minimo di ciascuna formante (F1 e F2) per le misurazioni effettuate sull'intero dittongo. Si ottengono così due misure distinte indicative dell'ampiezza delle transizioni delle prime due formanti. Tali misure sono rappresentabili in uno spazio F1-F2 come i cateti di un triangolo rettangolo. Pertanto, è possibile ottenere un unico indice dell'escursione dei valori delle due formanti, calcolando l'ipotenusa di questo triangolo attraverso il teorema di Pitagora. La formula per il calcolo dell'indice è dunque la seguente:

$$\text{coeff. ditt.} = \sqrt{(F1max - F1min)^2 + (F2max - F2min)^2}$$

Prima di essere immessi nella formula, i valori in Hertz vengono convertiti in Bark secondo la formula di Traunmüller (1990). In tal modo viene valutato in maniera più adeguata l'apporto dei movimenti della prima formante all'ampiezza totale del movimento dittongale. I movimenti della F1, infatti, risultano meno ampi se considerati dal punto di vista della scala acustica in Hertz, mentre vengono rappresentati più adeguatamente se considerati dal punto di vista della scala uditiva in Bark.

A tale indice è stato dato il nome di 'coefficiente di dittongazione'. Il coefficiente ha valore 0 solo nei casi di segmenti vocalici più brevi di 60 ms., giacché questi vengono caratterizzati da una sola misurazione delle formanti in un punto medio. Per quanto riguarda invece vocali più lunghe di 60 ms., il coefficiente di dittongazione si mantiene in genere al disotto del valore 1 per gli esiti che percettivamente risultano di tipo monottongale, mentre supera il valore 1.8 per gli esiti che percettivamente risultano di tipo dittongale. All'ascolto impressionistico dei dati sembra che segmenti caratterizzati da un coefficiente di dittongazione inferiore a 1.8 non producano la percezione di un dittongo[12].

[12] Questa "soglia critica" per la percezione di dittongazione è stata definita su base impressionistica. In futuro si prevede di approfondire meglio tale questione attraverso lo sviluppo di adeguati test percettivi.

5 Dittonghi e monottonghi nello spazio F1-F2

Nella figura 1 si riportano gli esiti monottongali e gli esiti dittongali nello spazio F1-F2, relativamente alle variabili vocaliche analizzate per il corpus di Pozzuoli. Gli esiti monottongali sono rappresentati dai simboli in grigio e corrispondono alle medie di F1 e F2 delle misurazioni formantiche delle realizzazioni in posizione interna della variabile corrispondente.

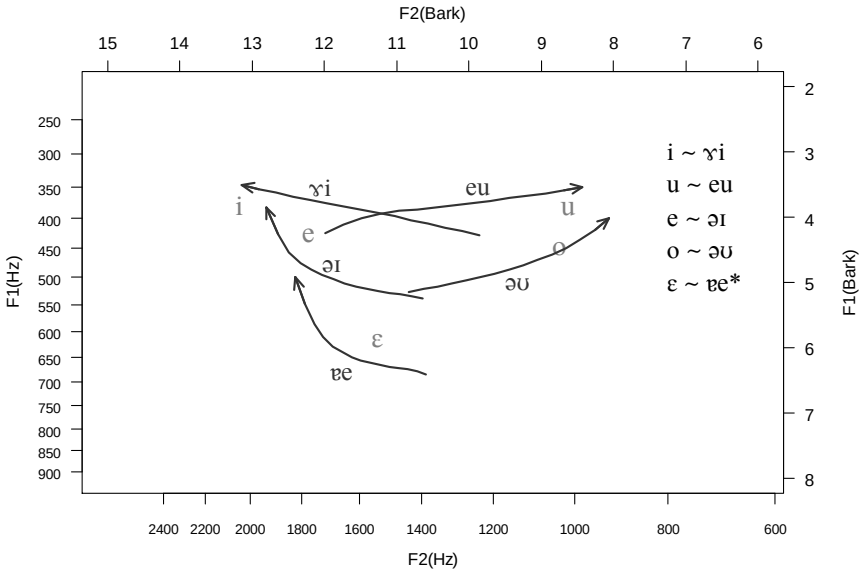


Figura 1: Traiettorie dittongali medie nello spazio vocalico

Per la resa delle traiettorie dittongali nello spazio vocalico si è proceduto in questo modo: sono stati presi i valori medi di F1 e F2 che caratterizzano le traiettorie dittongali per ogni classe di durata di una variabile[13] e per ciascuna classe sono stati calcolati cinque valori di F1 e F2 corrispondenti a: valore minimo, primo, secondo e terzo quartile, valore massimo; quindi è stata calcolata la media di questi cinque valori per tutte le classi di durata analizzate. È stata poi realizzata una

[13] In rapporto al metodo di caratterizzazione della dinamica dittongale qui utilizzato, i dittonghi possono essere distinti in classi di durata di 20 ms., di modo che i dittonghi di una stessa classe saranno descritti da uno stesso numero di misurazioni formantiche.

linea curva che interpola i valori F1-F2 così ottenuti. Tale linea, conclusa da una freccia che indica la direzione del movimento, costituisce dunque una stima della traiettoria dittongale media di una variabile nello spazio vocalico.
Nella legenda sono riportate le varianti monottongali di ciascuna variabile analizzata, accompagnate dalle corrispettive varianti dittongali. Si badi che la variante dittongale di (ɛ) è seguita da un asterisco, per sottolineare che tale variante caratterizza solo il parlato di una parte degli informatori, ossia i pescatori sotto i quarant'anni.
Il grafico evidenzia una particolare simmetria tra gli esiti dittongali delle vocali anteriori e posteriori, come tra gli esiti dittongali delle due vocali anteriori medie. In particolare, i dittonghi di (i) e (u) risultano quasi perfettamente speculari, mentre il dittongo di (ɛ), che costituisce un'innovazione dei parlanti più giovani, sembra modellarsi sul dittongo di (e), pur collocandosi in un'area più bassa dello spazio vocalico. Tutti i dittonghi analizzati sono caratterizzati da un movimento verso la periferia alta del trapezio: gli esiti dittongali delle vocali anteriori tendono a [i], quelli delle vocali posteriori tendono a [u]. Inoltre, mentre per gli esiti dittongali delle vocali medie il movimento parte dalla parte centrale dello spazio vocalico, per gli esiti di (i) e (u) si hanno movimenti più ampi, che partono dall'area posteriore medio-alta nel caso di (i), e da quella anteriore medio-alta nel caso di (u).

6 Giovani vs anziani: il caso di (ɛ)

Mentre relativamente alle altre variabili vocaliche analizzate il campione presenta caratteristiche piuttosto omogenee, per quanto riguarda la variabile (ɛ) i dati in nostro possesso ci consentono di distinguere, come accennato, tra un gruppo dei "giovani" di età inferiore ai quarant'anni, e un gruppo degli "anziani" di età superiore ai quarant'anni. L'opposizione è dovuta alla presenza nel gruppo dei giovani di varianti dittongali in posizione finale di sintagma intonativo per la variabile (ɛ), mentre tali varianti dittongali mancano quasi del tutto nel gruppo degli anziani. I giovani, quindi, hanno esteso a questa variabile il *pattern* di alternanza monottongo/dittongo in rapporto alla posizione prosodica, che già caratterizzava le vocali alte e medio-alte anteriori e posteriori.

L'estensione del fenomeno di dittongazione alla variabile (ɛ) nel gruppo dei giovani è stata valutata in maniera quantitativa, attraverso lo studio del coefficiente di dittongazione nelle diverse posizioni prosodiche definite, mantenendo distinti i due gruppi generazionali.
Nella figura 2 si riportano in un grafico del tipo *box plot* i dati relativi al coefficiente di dittongazione dei *tokens* di (ɛ), per le tre posizioni prosodiche definite per questa ricerca. Due diverse gradazioni di grigio distinguono i dati relativi al gruppo dei giovani da quelli relativi al gruppo degli anziani. La linea tratteggiata indica una soglia critica per la percezione di dittongazione, che è stata definita per il momento su base impressionistica. Quando i *tokens* hanno un coefficiente di dittongazione che superi la soglia di 1.8, l'impressione uditiva è quella di un dittongo. Il grafico mostra come in posizione interna al sintagma intonativo i dati del coefficiente di dittongazione siano sostanzialmente omogenei nei due gruppi[14]. In posizione finale, invece, si manifesta una chiara demarcazione tra il gruppo dei giovani, per i quali i coefficienti di dittongazione superano in maniera netta la soglia critica di 1.8, e il gruppo degli anziani, che si mantengono quasi completamente al disotto della soglia, con sporadiche realizzazioni di tipo dittongale[15].

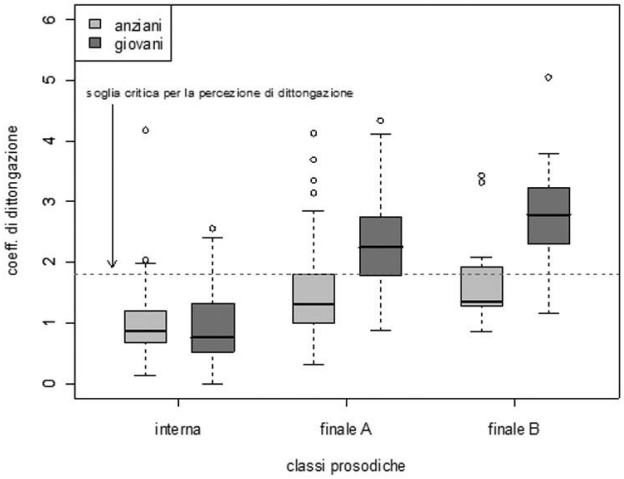


Figura 2: Coefficiente di dittongazione di (ɛ) per posizione prosodica tra giovani e anziani

[14] Come si vedrà, le differenze tra giovani e anziani in questa posizione non sono statisticamente significative.
[15] Si tratta in gran parte dei *tokens* relativi a toponimi, dei quali si discuterà in § 7.

Il grafico evidenzia anche come il coefficiente di dittongazione nel gruppo dei giovani presenti valori sensibilmente più alti nella posizione *Finale B*, caratterizzata da intonazione finale ascendente, rispetto alla posizione *Finale A*, priva di questo tipo di contorno intonativo. Anche per il gruppo degli anziani si manifesta un certo aumento del coefficiente di dittongazione nelle posizioni *Finale A* e *Finale B*, ma in tal caso si tratta di un aumento molto meno cospicuo, che non produce effetti uditivi significativi, ed è da ricondurre probabilmente alla maggiore durata che i segmenti vocalici presentano in queste posizioni anche nel gruppo degli anziani (cfr. figura 3).
La significatività statistica delle differenze notate nelle diverse classi prosodiche, come nei due gruppi generazionali, è stata valutata attraverso il test di Kruskal-Wallis (*One-way ANOVA by Ranks*)[16]. Per quanto riguarda l'opposizione tra gruppo dei giovani e gruppo degli anziani, le differenze non risultano statisticamente significative in posizione interna ($p. > 0.05$), mentre sono significative nelle posizioni finali di sintagma intonativo ($p. < 0.01$)[17]. Relativamente invece alle differenze che ciascun gruppo presenta nelle diverse classi prosodiche, sono statisticamente significative tutte le differenze notate, ad eccezione delle differenze nel gruppo degli anziani tra le classi *Finale A* e *Finale B* ($p. > 0.05$).
Ricapitolando, in posizione interna al sintagma intonativo giovani e anziani mantengono realizzazioni di (ε) sostanzialmente simili e di tipo monottongale; in posizione finale di sintagma intonativo i giovani presentano un notevole incremento dei valori medi del coefficiente di dittongazione (del 148% in *Finale A* e del 202% in *Finale B*), superando nettamente la soglia critica per la percezione di dittongazione, mentre per gli anziani l'incremento dei valori del coefficiente di dittongazione in tale posizione è molto più modesto (del 45% in *Finale A* e del 48% in *Finale B*), anche se statisticamente significativo, e i *tokens* presentano valori generalmente al disotto della soglia critica, quindi con esiti monottongali anche in posizione finale.

[16] Tale test costituisce un'alternativa non parametrica all'ANOVA. Per la sua definizione si rinvia a Hollander and Wolfe (1973: 115-20). La scelta di un test non parametrico è stata dettata dalla natura dei nostri dati, che non soddisfano alcuni requisiti delle statistiche parametriche, come la normalità delle distribuzioni e l'omogeneità delle varianze.
[17] Il risultato del test applicato all'insieme dei campioni (preliminarmente ai confronti appaiati) è: $\chi^2 = 145.2512$, df = 5, p. < 2.2e-16.

Di un certo interesse anche i dati relativi alle variazioni di durata nelle diverse classi prosodiche. Nella figura 3 si mostrano le durate dei *tokens* della variabile (ε) nelle tre classi prosodiche definite, mantenendo distinti i dati relativi al gruppo dei giovani e quelli relativi al gruppo degli anziani. In questo caso si rileva per entrambi i gruppi generazionali un analogo *pattern* di allungamento prepausale, anche se nei giovani l'allungamento è un po' più accentuato[18]. Questi dati mostrano come l'allungamento non sia una condizione sufficiente per l'emersione delle varianti dittongali[19]. I tokens di (ε) presentano lo stesso *pattern* di allungamento in rapporto alla posizione sia nei giovani che negli anziani, ma solo nel gruppo dei giovani all'allungamento si accompagna anche la dittongazione.

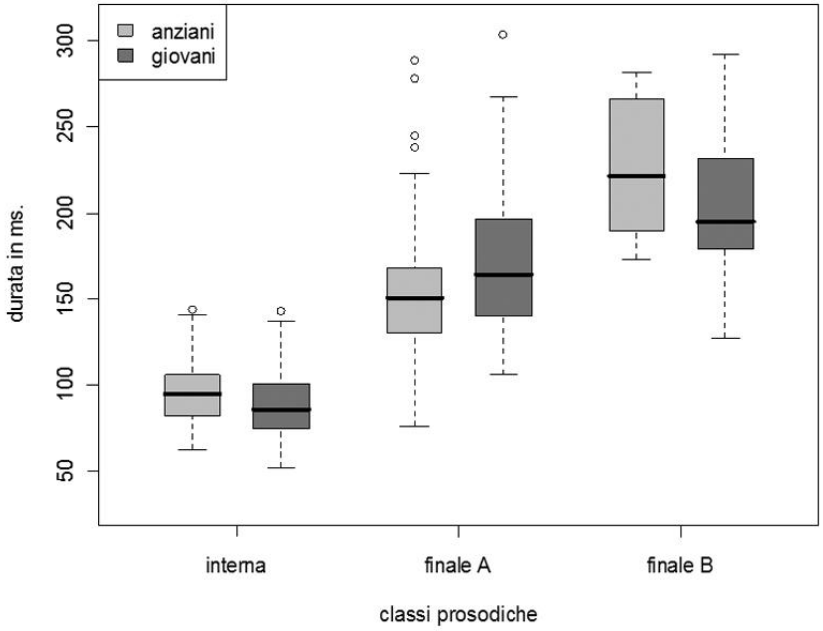


Figura 3: Durate di (ε) per posizione prosodica tra giovani e anziani

[18] Negli anziani l'incremento di durata dalla posizione interna alla posizione *Finale A* è del 56%, l'incremento dalla posizione interna alla posizione *Finale B* è del 127%; nei giovani l'incremento di durata dalla posizione interna alla posizione *Finale A* è del 78%, l'incremento dalla posizione interna alla posizione *Finale B* è del 109%.

[19] Sul rapporto tra posizione, durata e dittongazione cfr. Abete (in preparazione).

7 **Classi di parole**

I dati che abbiamo mostrato sembrano rientrare coerentemente nella casistica di un cambiamento in tempo apparente. I parlanti sotto i quarant'anni hanno esteso il *pattern* di variazione monottongo/ dittongo alla variabile (ɛ), che non è coinvolta da tale *pattern* nel parlato dei pescatori più anziani. Inoltre, il cambiamento sembrerebbe essere stato piuttosto veloce, in quanto nei parlanti sopra i quarant'anni le realizzazioni di (ɛ) sono quasi esclusivamente monottongali, mentre nei parlanti sotto i quarant'anni la dittongazione di (ɛ) in posizione finale si presenta come sistematica ed estesa a tutto il lessico[20]. Guardando alla diffusione del fenomeno nei giovani, alla sua sistematicità e regolarità, sembrerebbe di essere di fronte a un «regular sound change», secondo l'accezione di Labov (1994: 542), «without lexical or grammatical conditioning» (*ibid*.). Tuttavia, ad un'analisi più attenta dei dati, emergono indizi che portano a supporre che il cambiamento sia avvenuto piuttosto per diffusione lessicale, partendo da alcune classi di parole, ed estendendosi in seguito a tutti i lessemi comprendenti la variabile (ɛ).

A questo proposito, si riportano in figura 4 i dati del coefficiente di dittongazione in posizione finale per alcune classi di parole, distinguendo i valori del gruppo dei giovani da quelli del gruppo degli anziani. Se da un lato questo grafico conferma la netta distinzione tra i due gruppi generazionali, dall'altro si evidenzia anche come classi di parole diverse possano essere caratterizzate da coefficienti anch'essi piuttosto diversi[21]. Parte di queste differenze potrebbero dipendere plausibilmente dall'intorno fonetico. Sembra ad esempio che le varianti precedute da consonante bilabiale abbiano un coefficiente di dittongazione più alto, ma i dati ricavati dal parlato spontaneo non sono sufficienti a testare questa ipotesi con il dovuto rigore.

[20] Non si esclude comunque che un campione più consistente e meglio distribuito per fasce d'età potrebbe evidenziare un cambiamento meno discontinuo di quello qui rilevato.

[21] I dati relativi alle diverse classi di parole non sono molto consistenti e scoraggiano l'uso di test statistici (cfr. tab. 3). Ad ogni modo, l'impressione che la dittongazione di (ɛ) possa essere di diversa entità nelle varie classi di parole è confermato anche dall'osservazione dei grafici sull'andamento temporale dei valori formantici (cfr. Abete and Simpson, in press: § 4), che non è possibile però presentare in questa sede.

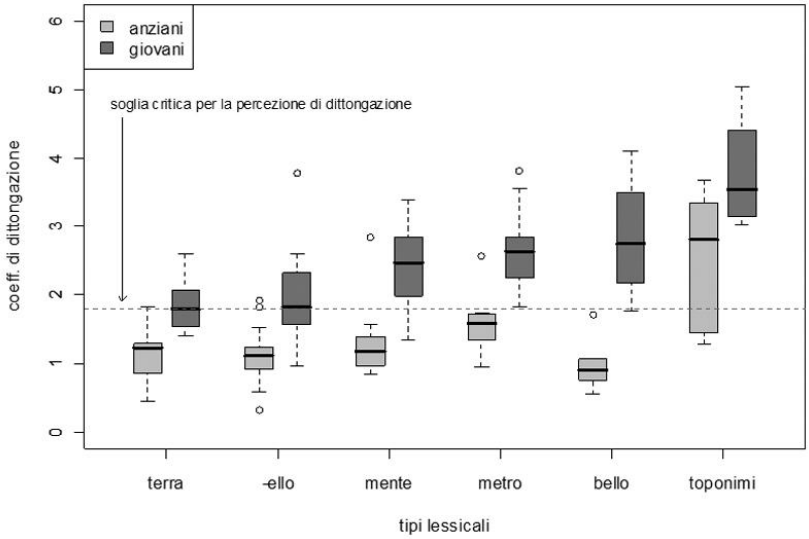


Figura 4: Coefficienti di dittongazione di (ε) in posizione finale, distinti per classi di parole

Il dato di maggiore interesse qui riguarda la classe dei toponimi, per la quale gli stessi parlanti anziani presentano coefficienti di dittongazione piuttosto alti, che superano spesso la soglia critica di 1.8. Si tratta di una classe piuttosto eterogenea dal punto di vista fonetico, che accoglie le realizzazioni di 5 nomi di luogo, *Civitavecchia*, *Maddalena*, *Marinella*, *Port'Ercole*, *Ventotene*, accomunati dal presentare la variabile (ε) nella sillaba tonica[22]. Di questa classe si hanno solo 4 realizzazioni nel gruppo dei giovani e 15 realizzazioni nel gruppo degli anziani. Per queste ultime, che sono qui di particolare interesse, si hanno 10 realizzazioni di *Civitavecchia*, 2 realizzazioni di *Marinella*, 3 realizzazioni di *Ventotene*. Negli anziani *Civitavecchia* presenta dittongazione regolarmente con coefficienti intorno a 3, una realizzazione di *Marinella* presenta un coefficiente di 2.08, mentre l'altra presenta un coefficiente di 1.55, le realizzazioni di *Ventotene* hanno coefficienti intorno a 1.3.

I dati a disposizione sono suggestivi, ma consentono per il momento solo riflessioni altamente speculative. È possibile che la dittongazione di (ε) sia partita da classi di parole come i toponimi, restando ad esse

[22] Tali toponimi non sono stati elicitati in maniera voluta, ma sono comparsi regolarmente nelle registrazioni, in quanto uno dei temi di conversazione adottati nelle interviste è stato quello dell'emigrazione stagionale in altri porti italiani.

limitata nel dialetto degli anziani, e si sia estesa invece a tutti i lessemi con (ɛ) nel dialetto dei giovani? In tal caso, l'estensione della dittongazione ad (ɛ) potrebbe rientrare nella casistica dei cambiamenti per diffusione lessicale che coinvolgono prima le parole meno frequenti, discussa da Phillips (2006). L'alternanza monottongo/ dittongo può essere intesa infatti come uno schema di variazione molto produttivo che caratterizza gran parte del vocalismo puteolano e tende ad estendersi anche ad altre variabili vocaliche, a scapito di forme tradizionali prive di varianti dittongali. In quest'ottica, tenderebbero ad assimilarsi a questo modello di variazione prima le forme meno frequenti e più soggette a fenomeni di livellamento analogico. Potrebbe essere questo il caso di parole come *Civitavecchia*? Il fenomeno è limitato a questa parola, o riguarda anche altri tipi lessicali che non sono emersi nelle interviste? Per rispondere a queste domande bisognerà tornare sul campo con tecniche di raccolta più intrusive di quelle qui utilizzate, e mirate all'elicitazione di speciali settori del lessico. Ad ogni modo, però, è l'osservazione di parlato spontaneo in situazioni comunicative reali che ha reso possibile la formulazione di tali quesiti.

8 Discussione

L'estensione della dittongazione nel dialetto dei giovani pescatori di Pozzuoli fornisce interessanti spunti di riflessione sulle dinamiche sociolinguistiche in atto in questa comunità. Come accennato nell'introduzione, il dialetto Puteolano è tra i più stigmatizzati in Campania. Nella numerose interviste raccolte nell'archivio ADICA (v. Sornicola, 2002) e relative alla provincia di Napoli, il dialetto puteolano viene molto frequentemente citato dai parlanti intervistati come dialetto "sguaiato" per antonomasia. Anche se i giudizi si soffermano raramente su caratteristiche fonetiche ben definite, è possibile dedurre da diversi indizi che tali opinioni si basino soprattutto sui fenomeni di dittongazione.
La stigmatizzazione non proviene solo dall'esterno e sono spesso gli stessi abitanti di Pozzuoli ad esprimere giudizi negativi nei confronti del proprio dialetto. Anche i pescatori intervistati, che costituiscono il gruppo sociale nel quale forse meglio si conserva il dialetto locale, hanno spesso atteggiamenti denigratori nei confronti della propria

parlata[23]. Inoltre, nelle produzioni concrete, le varianti dittongali si mostrano molto sensibili alle variazioni in diafasia, e vengono in genere rimosse e sostituite da varianti monottongali più vicine al modello napoletano, non appena la situazione comunicativa aumenti di formalità e il parlato si faccia più controllato[24].
È dunque possibile rintracciare, negli stessi parlanti intervistati per questa ricerca, giudizi e atteggiamenti[25] negativi di fondo sulla propria parlata e in particolare sulle varianti dittongali. Ciononostante, la dittongazione resta molto vitale nella comunità dei pescatori e, come si è visto in § 6, essa si è addirittura estesa a nuovi contesti nelle generazioni più giovani. I dati presentati mettono quindi in luce il complesso rapporto che sussiste tra giudizi, atteggiamenti e produzioni linguistiche concrete.
In una delle sue ricerche su Norwich, Trudgill (1972) fornisce evidenze empiriche sulla discrepanza tra le auto-valutazioni dei parlanti sulla propria parlata e i loro comportamenti effettivi, argomentando in favore di forme di "covert prestige" che, in contrasto con i giudizi espliciti, orienterebbero i comportamenti linguistici nelle situazioni più informali e spontanee. Similmente, in rapporto alla questione del perché le varietà vernacolari continuino ad esistere nonostante i giudizi negativi nei loro confronti, Labov postula «the existence of an opposing set of covert norms, wich attribute positive values to the vernacular» (Labov, 1972: 249)[26]. Tra i fattori "nascosti", che orienterebbero i comportamenti linguistici concreti, pur mantenendosi in genere sotto la soglia di consapevolezza, emerge in bibliografia soprattutto la solidarietà di gruppo, ossia l'esigenza dei parlanti di segnalare linguisticamente la propria appartenenza a un gruppo definito. Attraverso i risultati di test valutativi, Chambers (2009: 240) mette in evidenza una serie di qualità positive associate alle varietà sub-standard, riconducibili alla funzione di "solidarity stressing" (cfr. anche Chambers, 2000).
Il discorso relativo alla solidarietà di gruppo sembra particolarmente rilevante anche nell'interpretazione dei dati presentati in questa

[23] Per esempi concreti si rinvia ad Abete (2008).
[24] Alcune evidenze empiriche su questo aspetto sono presentate in Abete (2006: 160-168).
[25] Sul rapporto tra giudizi e atteggiamenti linguistici cfr. Berruto (1995: § 3.2.6).
[26] Cfr. anche Ryan (1979).

ricerca. Il gruppo dei pescatori intervistati è tenuto insieme da forti vincoli di solidarietà. Queste persone condividono gli stessi spazi, sono in parte imparentate tra loro, si aiutano spesso nel lavoro, discutono insieme i propri problemi e le decisioni da prendere. Tali vincoli risultano di particolare importanza, ad esempio, nei frequenti periodi di crisi economica, quando i pescatori si aiutano a vicenda. L'identità interna al gruppo si definisce anche in rapporto alle pressioni che provengono dall'esterno. Tale identità è infatti minacciata da fattori di disgregazione, che negli ultimi anni hanno agito in maniera molto forte. Dopo i moti bradisismici degli inizi degli anni '70, il centro storico di Pozzuoli, sino ad allora depositario del dialetto e della tradizione locale, è stato completamente evacuato. La sua popolazione è stata dislocata in varie aree periferiche e costretta a modificare le proprie abitudini di vita, i rapporti col vicinato, le modalità di lavoro[27]. Oggi i pescatori sono gli unici a condividere ancora uno spazio comune tradizionale, la zona del porto, dove ogni giorno si recano in auto per lavorare. Per molti aspetti, dunque, queste persone costituiscono l'ultimo baluardo di una società che va scomparendo. È possibile che queste dinamiche rafforzino ancora di più l'esigenza di coesione interna al gruppo. In quest'ottica, l'estensione della dittongazione nel dialetto dei giovani pescatori si carica di possibili valori sociolinguistici, in quanto va a rafforzare proprio uno di quei tratti che sono maggiormente distintivi della cultura locale e in particolare del gruppo dei pescatori. Queste osservazioni sono comunque di carattere solo preliminare e andranno approfondite attraverso ulteriori studi, in particolare allargando il corpus ad altri parlanti e ad altri gruppi sociali e applicando specifiche metodologie di tipo percezionale. Per il momento ci sembra di poter adattare alla comunità dei pescatori di Pozzuoli le parole scritte da Trudgill a proposito degli informatori di Norwich:

> [...] statements about "bad speech" are for public consumption only. Privately and subconsciously, a large number of male speakers are more concerned with acquiring prestige of the covert sort and with signalling group solidarity than with acquiring social status, as this is more usually defined (Trudgill, 1983: 177).

[27] Cfr. Giglia (1997).

Riferimenti bibliografici

Abete, G. (2006) 'Il polimorfismo delle realizzazioni vocaliche nel dialetto di Pozzuoli'. *Bollettino Linguistico Campano*, 9/10, pp. 143-72.

– (2008) Lo status sociolinguistico della dittongazione spontanea. Realtà campane e calabresi a confronto, in G. Marcato (ed.) *L'Italia dei dialetti*, Atti del Convegno di Sappada/Plodn (Belluno), 27 giugno-1 luglio 2007, pp. 307-14.

– (in preparazione) *I processi di dittongazione nei dialetti dell'Italia meridionale. Un approccio sperimentale*, Tesi di dottorato, Friedrich-Schiller-Universität, Jena.

Abete, G. and A.P. Simpson (in press) Confini prosodici e variazione segmentale. Analisi acustica dell'alternanza monottongo/ dittongo in alcuni dialetti dell'Italia meridionale, in *Atti del 5° Convegno Nazionale AISV*, Zurigo, 4-6 febbraio 2009.

Berruto, G. (1995) *Fondamenti di sociolinguistica*. Bari: Laterza.

Chambers, J.K. (2000) Sociolinguistic uses of subjective evaluation texts, in S. Deninger, T. Förgen, J. Scharloth and S. Zwickl (eds) *Einstellungsforschung in der Soziolinguistik und Nachbardisziplinen*, Frankfurt: Peter Lang, pp. 73-81.

– (2009) *Sociolinguistic theory*. Revised edition. Oxford: Wiley-Blackwell.

Como, P. (2006) Elicitation techniques for spoken discourse, in K. Brown (ed.) *Encyclopaedia of language and linguistics*, second edition, vol. 4, Amsterdam: Elsevier, pp. 105-09.

Giglia, A. (1997) *Crisi e ricostruzione di uno spazio urbano dopo il bradisismo a Pozzuoli: una ricerca antropologia su Monte Ruscello*. Milano: Guerini.

Holbrook, A. and G. Fairbanks (1962) 'Diphthong formants and their movements'. *Journal of Speech and Hearing Research*, 5, pp. 38-58.

Hollander, M. and D.A. Wolfe (1973) *Nonparametric statistical methods*. New York: John Wiley & Sons.

Labov, W. (1972) *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press.

– (1994) *Principles of linguistic change*. Vol. 1: *Internal factors*. Oxford: Blackwell.

Nespor, M. (1993) *Fonologia*. Bologna: il Mulino.
Phillips, B.S. (2006) *Word frequency and lexical Diffusion*. New York: Palgrave Macmillan.
Rohlfs, G. (1966) [1949] *Grammatica storica della lingua italiana e dei suoi dialetti. 1: Fonetica*. Torino: Einaudi.
Ryan, E. B. (1979) Why do low-prestige language varieties persist?, in H. Giles and R. St Clair (eds) *Language and Social Psychology*, Oxford: Basil Blackwell, pp. 145-57.
Simpson, A. P. (1998) Characterizing the formant movements of German diphthongs in spontaneous speech, in B. Schröder, W. Lenders, W. Hess and T. Portele (eds) *Computer Linguistik und Phonetik zwischen Sprache und Sprechen*, Tagungsband der 4. Konferenz zur Verarbeitung natürlicher Sprache – KONVENS – 98, Frankfurt: Lang, pp. 192-200.
Sornicola, R. (2002) 'La variazione dialettale nell'area costiera napoletana. Il progetto di un archivio di testi dialettali parlati'. *Bollettino Linguistico Campano*, 1, pp. 131-55.
Traunmüller, H. (1990) 'Analytical expressions for the tonotopic sensory scale'. *Journal of the Acoustical Society of America*, 88, pp. 97-100.
Trudgill, P. (1972) 'Sex, covert prestige, and linguistic change in the urban British English of Norwich'. *Language in Society*,1, pp. 179-96.
– (1983) *On dialect: social and geographic factors*. Oxford: Blackwell.

Analisi prosodica della cronaca in diretto del giornalista sportivo

Georges Boulakia, Catherine Mathon, Takeki Kamiyama***
EA333 ARP, UFRL case 7003, Université Paris Diderot – Paris 7
*Université Bordeaux 3, UFR Lettres, Dpt Sciences du langage
**ILPGA, Université de la Sorbonne Nouvelle – Paris 3
georges.boulakia@linguist.jussieu.fr; catherine.mathon@u-bordeaux3.fr;
takeki.kamiyama@univ-paris3.fr

Abstract

Sport commentary has two different aspecs. First it is shared by several speakers, each of whom plays a precise role throughout the match. Second, this commentary is made up of different types of discourse sequences. On the one hand there are 'descriptive' sequences in which the reporter describes what happens in front of his own eyes; this description is motivated by the action going on in real time and it leads us to think that these play-by-play actions will be supported by prosodic variations. On the other hand, these descriptive sequences alternate with either pseudo-monologue (the reporter knows that he is actually addressing an audience) or pseudo-dialogue sequences when the commentator develops technical or strategic considerations during the game (for example when the expert makes comments). Our hypothesis is that these types of discourse are characterised by different prosodic features, as well as by syntactic constructions. Our corpus consists of live TV commentaries recorded during the last (2007) Rugby World Cup. We recorded 36 hours of TV programmes which permit a 'multi-modal' (sounds and images) analysis. We selected the games in which commentaries are available in French, English and Japanese and these three languages are analysed. We analyse variations of melodic patterns as well as sudden changes of tempo or rhythm according to the action actually seen on the screen.

1. Premessa

La cronaca in diretta del giornalista sportivo presenta molteplici specificità linguistiche che ne fanno un genere di discorso a sè e, in quanto genere proprio, facilmente identificato dagli ascoltatori. D'altronde, e ciò sarà oggetto di una delle parti della nostra analisi (v. 4.2.), sembra che tali specificità linguistiche (lessicali, sintattiche, prosodiche) siano comuni a diverse lingue. Nel presente articolo, ci interesseremo più particolarmente delle specificità prosodiche e

mostreremo che esse sono comuni al francese, all'inglese e al giapponese. Se, malgrado le rilevanti differenze prosodiche esistenti fra queste tre lingue, si riscontrano dei caratteri comuni nel discorso del cronista sportivo, ciò è probabilmente dovuto al fatto che alle strutture prosodiche organizzative inerenti a ciascuna lingua si sovrappone un livello superiore che costituisce lo stile prosodico di questo genere di discorso. Questo stile prosodico è legato in parte a considerazioni pragmatiche che tenteremo di mettere in luce.

1.1. Iconicità della parola

La cronaca sportiva presenta un'altra particolarità interessante che si riflette direttamente a livello prosodico: siccome una parte del commento consiste nella descrizione in diretta di un'azione che si svolge sotto gli occhi del locutore, il discorso del cronista sportivo è motivato dall'azione in corso: il discorso si costruisce contemporaneamente all'azione. Esiste una sorta di sincronicità fra il tempo dell'enunciazione e il tempo dell'azione di gioco. Questo legame sincronico fra il discorso descrittivo e l'azione di gioco si ripercuoterà sulla prosodia del locutore. Nella sezione 4.2. mostreremo che il ritmo della parola del locutore corrisponde al ritmo dell'azione sul terreno di gioco.

È risaputo che il segno linguistico è arbitrario e non intendiamo rimettere in questione questo dato acquisito della linguistica descrittiva da Saussure in poi. Tuttavia, non si possono non sottolineare a livello prosodico taluni fenomeni che sembrano riflettere l'azione in corso. E, in fin dei conti, niente di più normale, nella misura in cui il locutore cerca, attraverso il suo discorso descrittivo, di essere il più possibile "aderente" all'azione che si svolge sotto il suo sguardo affinché i suoi ascoltatori si rappresentino nel migliore dei modi la scena descritta e ne condividano l'emozione. È indispensabile che l'ascoltatore abbia l'impressione di assistere alla partita come se fosse in una tribuna dello stadio. Attraverso queste specie di metafore vocali (come la caduta repentina della curva melodica che coincide con l'uscita non meno repentina del pallone dallo specchio della porta), il locutore cerca soltanto di far condividere al suo pubblico, nel modo più diretto possibile, lo spettacolo al quale assiste e di renderlo più vivo e più presente per il telespettatore.

1.2. Strutturazione sintattica del commento sportivo

Per terminare con le peculiarità del commento sportivo, vorremmo insistere sulla sua strutturazione linguistica e in particolarre sintattica, nella misura in cui essa scaturisce dalla situazione enunciativa di questo discorso e dalla sua iconicità, a sua volta indotta dalla situazione enunciativa. Metteremo l'accento sulle particolarità sintattiche del discorso descrittivo, vale a dire della cronaca del giornalista sportivo, motivata e sincronica con l'azione di gioco.
Il discorso descrittivo della cronica sportiva presenta la particolarità di essere essenzialmente costituito da costruzioni nominali (Deulofeu, 1998). Per costruzioni nominali si intendono degli enunciati il cui nucleo non comporta sintagma verbale. In effetti, in questo discorso, si riscontra un gran numero di costruzioni nominali e, benché non siano proprie di questo genere, esse sono particolarmente frequenti. Queste costruzioni offrono il vantaggio di adattarsi facilmente agli imprevisti del discorso motivati dagli imprevisti dell'azione. In una prospettiva macro-sintattica (Blanche-Benveniste, 2000) associata alla teoria del contrasto delle *pentes di Martin (Martin, 1981; Boulakia *et al*, 2001), l'organizzazione e l'interpretazione di tali costruzioni nominali sono fornite dalla prosodia.

2. Ipotesi

In questo lavoro intendiamo esplorare i diversi aspetti della cronaca sportiva testè menzionati secondo tre ipotesi.
Il primo punto concerne la specificità della situazione di enunciazione, ed in particolare il fenomeno del tropo comunicativo. L'ipotesi è che la presenza virtuale di una terza persona nella situazione enunciativa e la sua presa in considerazione da parte del locutore avrà un'influenza non solo sul messaggio veicolato ma anche sul modo in cui il messaggio è veicolato e quindi sulla voce. Formuliamo dunque l'ipotesi che la presa in considerazione del telespettatore da parte del giornalista si tradurrà nel suo discorso attraverso l'utilizzazione quasi costante di un registro di voce alto.
Il secondo punto riguarda il discorso descrittivo del giornalista rispetto all'azione. Abbiamo già ricordato che il discorso del giornalista è motivato dall'incontro che si svolge sotto il suo sguardo. Orbene:

durante una partita, fasi di gioco intense e rapide si alternano a periodi più lenti; se il discorso del cronista è effettivamente motivato dalla partita, ci si potrebbe aspettare che questa successione di fasi lente e più rapide si traduca o nel contenuto lessicale del discorso o nella prosodia del discorso descrittivo, o in ambedue gli ambiti. A livello prosodico, si può presumere che i periodi di gioco più lenti saranno caratterizzati dal locutore con un registro più basso, mentre, per contrasto, i periodi di maggiore attività si tradurranno attraverso un registro vocale alto e una brusca accelerazione della velocità di eloquio.
Il terzo punto concerne in parte l'aspetto iconico di questo tipo di discorso. In effetti, se la cronaca sportiva presenta degli aspetti iconici, è perché essa è motivata dall'azione. La nostra ipotesi è la seguente: se il discorso è effettivamente iconico, questi aspetti appariranno in modo identico, qualunque sia la lingua nella quale è pronunciato il commento. In questa prospettiva, le caratteristiche prosodiche del commento sportivo sarebbero dunque le stesse in lingue tanto diverse quali il francese, l'inglese e il giapponese. Discuteremo in 5. se queste similitudini da una lingua all'altra attengono al carattere iconico di questo tipo di linguaggio oppure ad uno stile giornalistico ampiamente diffuso dai media.

3. Corpus

Questo studio si inserisce in un progetto più vasto basato su un corpus esteso di commenti sportivi registrati nel corso della Coppa del Mondo di rugby 2007. Le registrazioni sono state effettuate dal 7 settembre al 20 ottobre. Da questo corpus, abbiamo scelto alcuni incontri che ci sono sembrati rappresentativi. Indicheremo quindi in un primo tempo l'origine e la costituzione di questo corpus e in un secondo tempo descriveremo il corpus di riferimento selezionato per questo studio.

3.1. Insieme del corpus

Il nostro corpus orale si compone di 2 sub-corpora corrispondenti a due tipi di locutore e di registrazione
A il corpus dei cronisti
B il corpus dei tifosi (spettatori)

Il nostro primo sub-corpus comprende la cronaca sportiva di giornalisti. Essi commentano in diretta gli incontri del campionato. Gli incontri sono stati digitalizzati in modo da mantenere l'audio, versione mono, con una frequenza di campionamento di 22050 Hz o 44100 Hz. Abbiamo registrato le cronache nelle rispettive lingue dalla televisione francese, inglese e giapponese. Abbiamo raccolto le registrazioni di 19 incontri per il francese, per un equivalente approssimativo di 36 ore di registrazione. Per le altre lingue, i corpora sono più limitati ma avevamo l'intenzione di trattarli in altro modo allo scopo di confontare i diversi comportamenti dal punto di vista culturale. Abbiamo registrato solo due incontri (corrispondenti a 3.3 ore) per l'inglese e tre incontri (corrispondenti a 5.6 ore) per il giapponese. Il corpus giornalistico è quindi composto da 24 incontri per un totale di 45.3 ore di registrazione.
Il secondo sub-corpus è diverso da quello giornalistico sia dal punto di vista del locutore che da quello del tipo di registrazione. Abbiamo registrato i commenti di tifosi che assistevano all'incontro davanti al televisore o davanti al grande schermo all'aperto. Questo corpus riguarda solo il francese. Abbiamo raccolto registrazioni sonore da 18 locutori, 14 voci maschili e 4 femminili, nel corso di 7 diversi incontri, per un totale di 10.55 ore di registrazione.

3.2. Corpus di riferimento

Per il nostro studio, abbiamo lavorato solo sul corpus di cronache sportive prodotte dai giornalisti. Siccome ci situiamo in una prospettiva contrastiva, al fine di verificare la nostra terza ipotesi abbiamo privilegiato gli incontri per i quali avevamo a disposizione il commento in almeno due lingue. D'altro canto, volendo verificare se l'implicazione del cronista nei confronti di una delle due squadre potesse influenzare il suo discorso, in particolare a livello prosodico, abbiamo cercato di lavorare su partite nelle quali almeno uno dei giornalisti era implicato a causa della sua nazionalità nei confronti dell'una o dell'altra formazione in campo. Abbiamo così selezionato quattro incontri.

Incontri	Squadre in campo	Lingua/e della cronaca
1	Francia-Argentina	Francese
2	Francia-Inghilterra	Francese/Inglese
3	Argentina-Scozia	Francese/Inglese
4	Giappone-Fidji	Francese/Giapponese

Tabella 1: Scelta degli incontri e lingue

3.3. Situazione enunciativa

La diretta sportiva offre una situazione di enunciazione interessantissima e inedita, della quale descriveremo le componenti.

3.3.1. Locutori

Si tratta di un discorso condotto da un locutore principale, il giornalista sportivo, che descrive e commenta l'incontro; torneremo più oltre su queste due funzioni. Questo primo locutore condivide in modo più o meno calcolato il suo tempo di parola con un secondo locutore, l'esperto, il quale ha come funzione principale quella di commentare le strategie di gioco ed eventualmente spiegare talune azioni specifiche. Il tempo di parola della cronaca è quindi condiviso dai due locutori e presenta un'alternanza di fasi di monologo nelle quali il giornalista sportivo è l'unico locutore – e vedremo più avanti che questo monologo corrisponde a un tipo di discorso particolare – e fasi di dialogo quando il giornalista sportivo cede la parola all'esperto. D'altronde, un terzo locutore, che chiameremo "il giornalista sul terreno di gioco", interverrà con tempi di parola molto brevi. Si tratta di un giornalista sportivo situato in prossimità immediata del terreno di gioco e che interverrà in modo molto puntuale per segnalare i cambi di giocatori o per commentare un'atmosfera oppure l'umore dei giocatori… Questo terzo locutore sarà appena preso in considerazione, tenuto conto della poca importanza dei suoi interventi, in termini quantitativi.

3.3.2. Monologo vs dialogo

Queste fasi di monologo e di dialogo si riallacciano, non a caso, alla terminologia teatrale: come in una pièce teatrale i ruoli di ciascun

locutore che costruisce il commento sportivo sono accuratamente distribuiti. Così, le fasi di dialogo sono attentamente orchestrate in modo da dare un carattere interattivo e conviviale alla cronaca. Il giornalista sportivo e l'esperto rappresentano una coppia di amici riunitisi per passare una buona serata in compagnia godendosi una bella partita. Ci soffermeremo in particolare sul tono confidenziale che il giornalista può assumere per interrogare il suo compagno (Figura 3). Malgrado ciò, questo dialogo non ha nulla di spontaneo: esso non obbedisce in senso proprio a una sceneggiatura accuratamente prestabilita visto che il discorso si costruisce parallelamente all'evolversi dell'incontro; tuttavia, i tempi di parola del locutore-esperto sono attentamente definiti e distribuiti dal locutore-giornalista. Queste fasi di pseudo-dialogo si alternano a fasi di monologo soprattutto da parte del cronista (anche se taluni interventi lunghi, specialmente dell'esperto, si possono considerare dei monologhi). Questi monologhi intervengono quando, sul terreno di gioco, l'azione in corso merita una descrizione ed hanno quindi generalmente la struttura di un discorso descrittivo. Si può infatti postulare che, durante un incontro, ci sia una successione di fasi di gioco attive o più lente. Quando si presentano delle fasi attive, il cronista passerà naturalmente da una fase di dialogo, che serviva in realtà a "riempire i vuoti", a una fase di monologo incentrata sulla descrizione dell'azione in corso. In questo senso, queste fasi di monologo descrittivo sono motivate dall'azione che si svolge sotto gli occhi del locutore. Per terminare con questa fase di monologo, aggiungeremo che si tratta in effetti di uno pseudo-monologo in quanto, come nella rappresentazione teatrale, il locutore non pronuncia questo monologo per se stesso, e neppure per colui che sarebbe, almeno in senso spaziale, il suo interlocutore più diretto, cioè il locutore-esperto. Quest'ultimo, del resto, non ne ha alcun bisogno visto che ha lo stesso punto di osservazione del giornalista. A teatro, le fasi di monologo servono ad informare lo spettatore a proposito di avvenimenti passati o futuri, a precisare gli aspetti rimasti oscuri della trama o gli stati d'animo di un personaggio… Nei monologhi del discorso sportivo, il cronista si rivolge in realtà allo spettatore per descrivere l'azione in corso.

3.3.3. Un interlocutore principale presente seppur invisibile

È forse una delle particolarità più notevoli di questa situazione enunciativa: il vero interlocutore, colui al quale si rivolge l'intero discorso prodotto dal cronista sportivo è in realtà un interlocutore virtuale, in quanto non presente concretamente nello spazio e nel tempo dell'enunciazione. Non si tratta neppure di un interlocutore chiaramente definito, siccome si tratta del pubblico dei telespettatori, vale a dire di una sorta di entità composta da migliaia di individui indeterminati, che agli occhi del locutore si riducono a parti di audience. Malgrado questa indistinzione del telespettatore, è quasi esclusivamente attorno a lui che si concentra tutto il discorso del cronista, il quale pur tuttavia fa in fondo poco riferimento ad esso: al massimo, si rivolge di tanto in tanto direttamente al telespettatore, soprattutto nelle fasi introduttive della cronaca. L'interlocutore, benché non menzionato (o molto poco) nel messaggio, è pur tuttavia preso in considerazione ad ogni istante dal locutore. Siamo quindi di fronte ad una situazione di tropo comunicativo (Kerbrat-Orecchioni, 1990), fenomeno abbastanza comune nelle situazioni di enunciazione comportanti l'impiego dei media come pure in situazioni di recita (a teatro in particolare). Questa presa in considerazione appare a livello non più lessicale ma prosodico. È ciò che vedremo nella parte 4.

4. Analisi

Presentiamo un'analisi essenzialmente prosodica. Faremo talora allusione ad aspetti semantici o sintattici, ma questi saranno sempre evocati in ragione dei loro rapporti con la struttura prosodica.
Inoltre, la nostra analisi concerne principalmente il francese. In effetti, il corpus più importante di cui disponiamo è in francese. Ci siamo dunque interessati, per cominciare, all'analisi della fase di commento del discorso del cronista sportivo, quella nella quale egli situa l'incontro, parla delle squadre, ricorda gli incontri precedenti… Gli esempi proposti sono tratti dall'incontro Francia-Argentina, nella parte della cronaca che precede l'inizio dell'incontro, in modo da caratterizzare un discorso non ancora impregnato dall'eccitazione agonistica.
Abbiamo in seguito analizzato la parte descrittiva del discorso, per mostrarne le specificità prosodiche e in particolare in quale modo

delle fasi attive succedono a fasi più tranquille e come questo concatenamento riflette il concatenameno delle azioni di gioco. Anche in questo caso, abbiamo basato la nostra analisi sul francese. Inoltre, in un'ottica contrastiva, abbiamo sottoposto ad analisi alcuni esempi dell'inglese e del giapponese. Capita ancora molto raramente di lavorare su corpus orali comparabili in lingue diverse poiché, molto spesso, si pone il problema della pertinenza della comparazione: come si può stabilire che due segmenti di enunciazione sono comparabili? Si possono considerare comparabili quando hanno almeno un punto in comune. Nel nostro caso, la situazione di enunciazione è la stessa. I due cronisti producono nelle due diverse lingue un discorso in presenza dello stesso evento, indirizzato ad un interlocutore invisibile, il telespettatore. Inoltre, una parte del discorso prodotto, quella descrittiva, è motivata dall'azione di gioco in corso di realizzazione, vale a dire che una stessa azione di gioco può produrre discorsi sufficientemente simili (dunque comparabili) e, questo, qualunque sia la lingua nella quale il discorso si realizza.
Prima di guardare nei dettagli alcuni enunciati prodotti nel corso degli incontri, sia in francese che in inglese e in giapponese, abbiamo tentato di caratterizzare in modo molto globale la voce del locutore principale. A tal fine, abbiamo lavorato sulla cronaca in francese dell'incontro Francia-Argentina. Questa registrazione è stata tradotta con il programma Transcriber, allineata e analizzata con il programma WinPitch Pro (di Philip Martin, http://www.winpitch.com/). Questo programma consente l'estrazione automatica dei valori della F0 sui turni di parola del locutore principale ogni 50ms. Abbiamo in seguito individuato i valori minimi e massimi della F0, valori che abbiamo verificato manualmente. Avendo ottenuto in questo modo la gamma della voce del locutore, l'abbiamo suddivisa in quattro registri uguali: i registri alto, medio-alto, medio-basso e basso. La tabella 2 dà i valori frontiera di ciascuno di questi registri per il cronista dell'incontro Francia-Argentina.

	450 Hz
HIGH	
	365 Hz
MEDIUM-HIGH	
	280 Hz
MEDIUM-LOW	
	195 Hz
LOW	
	110 Hz

Tabella 2: Registro vocale di un giornalista francese

In seguito, analizzando diversi enunciati prodotti da questo locutore e calcolando la sua F0 ci si rende conto molto in fretta che il locutore si situa in un registro vocale alto e, ciò, durante tutta la cronaca. Questo semplicissimo calcolo ci consente di avere dei valori di riferimento per analizzare in modo dettagliato le produzioni dello stesso locutore in diversi momenti del discorso.

4.1. Le cronache

Prenderemo quattro esempi in francese, esempi tipici nel senso che si ritrovano ovunque, ivi compreso nei discorsi in inglese o in giapponese.
Questi esempi sono analizzati e visualizzati grazie a WinPitch Pro.
Partendo dal basso, la prima linea rappresenta le variazioni di intensità espresse in decibel (dB) (scala delle ordinate); questo parametro non verrà preso in esame; si può constatare che il rumore di fondo va dai 20 ai 30 dB, ma il rapporto segnale/rumore resta accettabile visto che la parola può essere isolata da questo rumore fino a 40 dB (limite probabilmente imposto automaticamente dalla registrazione televisiva).
La seconda linea rappresenta l'onda sonora costituita dal segnale di parola dei locutori come pure dal rumore di fondo. Il segnale si sviluppa sull'asse temporale (orizzontale) calibrato in secondi. Esso consente di calcolare le variazioni di durata e di velocità di eloquio.
La terza linea rappresenta la variazione della melodia della parola espressa in Herz (Hz) estratta automaticamente dal programma (funzione "pettine" di Martin).
La quarta "linea" rappresenta lo spettrogramma dell'esempio sonoro analizzato. Esso è calibrato in centinaia di Hz (sull'asse delle ordinate); l'abbiamo lasciato in "filtraggio stretto" per mostrare che il movimento delle armoniche corrisponde correttamente all'estrazione della F0 automatica della linea 3. Questa precauzione è necessaria in presenza di registrazioni con rumori di fondo.
Queste rappresentazioni consentono quindi di oggettivare le variazioni dei parametri e di caratterizzare i registri. La lunghezza degli esempi non permette un allineamento suono/testo.
Nel primo esempio (figura 1) si tratta di una presentazione contrastiva di 2 elementi ("d'un côté, … de l'autre …"). Si osserva (e si sente!) un

duplice movimento melodico con un'ascesa iniziale sul primo termine seguita da una discesa della F0 sul secondo termine; ma l'apice non è raggiunto: si compone esso stesso di una discesa/ascesa, dove la discesa rappresenta la fine (parentesi) del primo termine (il nome di un giocatore) e l'ascesa l'annuncio del secondo termine (il nome di un altro giocatore) della presentazione ("..de l'autre..") (Si potrebbe descrivere la variazione della FO qui come un"'sinclinale montonato melodico")

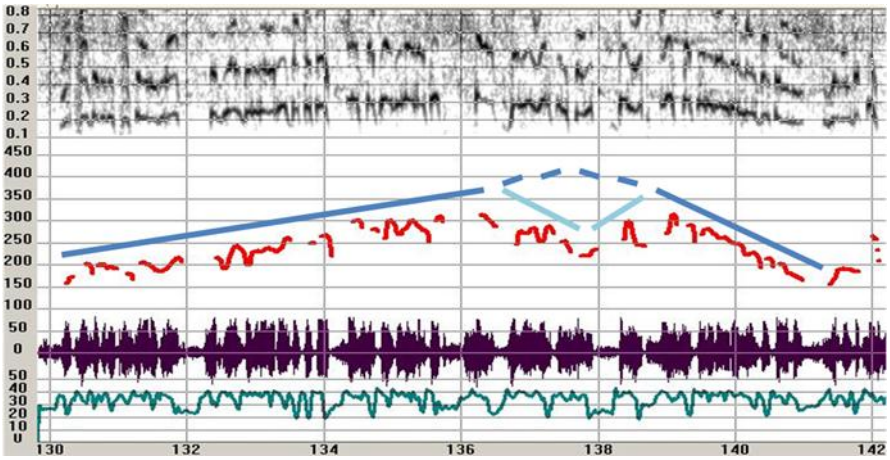


Figura 1: « La troisième illustration elle est évidente c'est le duel des numéros dix d'un côté un champion du monde en titre ↗ Sir Jonny Wilkinson ↘ # ↗ de l'autre ↘ un champion du monde des moins de vingt et un an Lionel Beauxis »

Nel secondo esempio, si tratta della presentazione di un celebre giocatore. All'ascolto, l'impressione è quella di un'enumerazione a causa della prominenza (picchi della curva melodica) della prima sillaba della parola, dei pronomi "ils", degli articoli, di "et". Si nota un'alternanza di registri medio/elevato.

Si tratta di una strategia deliberata che si ritrova in altri momenti, e nelle tre lingue, allorquando il giornalista presenta un giocatore, un avvenimento noto. Il giornalista si rivolge agli ascoltatori/spettatori invisibili, tenendoli col fiato sospeso: si tratta della funzione di appello.

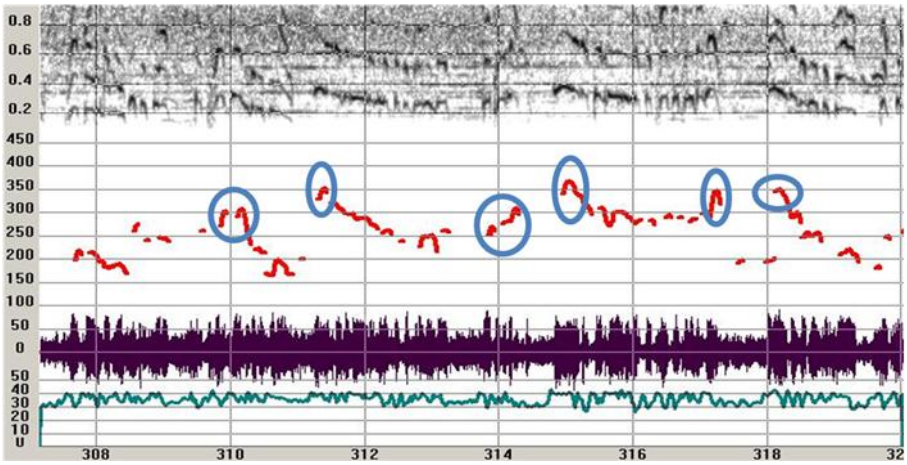


Figura 2: "« Raphaël Ibanez joue ce soir sa troisième demi-finale il en a gagné une en mille neuf cent quatre-vingt dix-neuf (↘) contre les Blacks il en a perdu une il y a quatre ans contre les Anglais et il veut gagner ce soir. »

Nei due esempi seguenti, si tratta di pseudo-dialoghi fra due "complici". Nella figura 3, si tratta di una domanda in aparté: registro basso e melodia poco mossa (piatta) con ascesa di continuità su "pelouse" (= interruzione per riprendere fiato) prima di continuare, sullo stesso registro basso, con il seguito della domanda retorica.

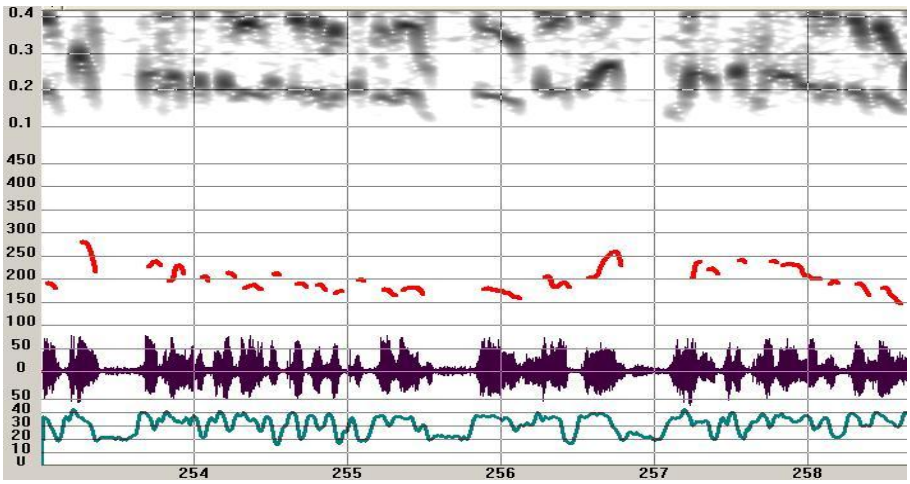


Figura 3: tiens Fabrice et Thierry avant que les deux équipes n'arrivent euh sur la pelouse (↗) # j'ai un petit problème à vous poser

Nella figura 4, troviamo un esempio tipico di turno di parola; il primo locutore realizza una caduta finale di F0 su “Fabrice” (il nome del collega complice) per passare la parola al secondo locutore che risponde con una ascesa iniziale.

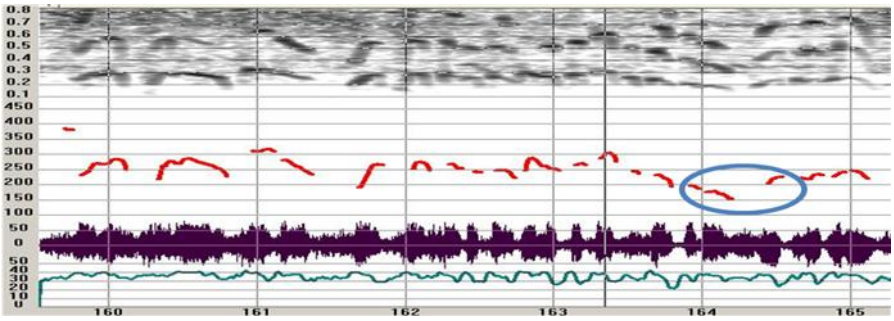


Figura 4: en Australie il pleuvait les Français avaient perdu il ne pleut pas ce soir Fabrice (↘) # (↗) non tout va bien.

Per quanto riguarda i giornalisti inglesi e giapponesi, abbiamo rilevato lo stesso tipo fonostilistico, simile dunque a quello del francese, ma naturalmente con delle differenze di ritmo, d’accentuazione, legate alle caratteristiche prosodiche proprie di queste due lingue.

4.2. La descrizione in diretta e i suoi vincoli

Partiremo da un esempio in francese che illustra bene il carattere iconico della prosodia: seguendo la linea melodica si ha l’impressione di seguire la fase di gioco (accelerazione e rallentamento) e soprattutto la traiettoria del pallone (che sale e ridiscende).

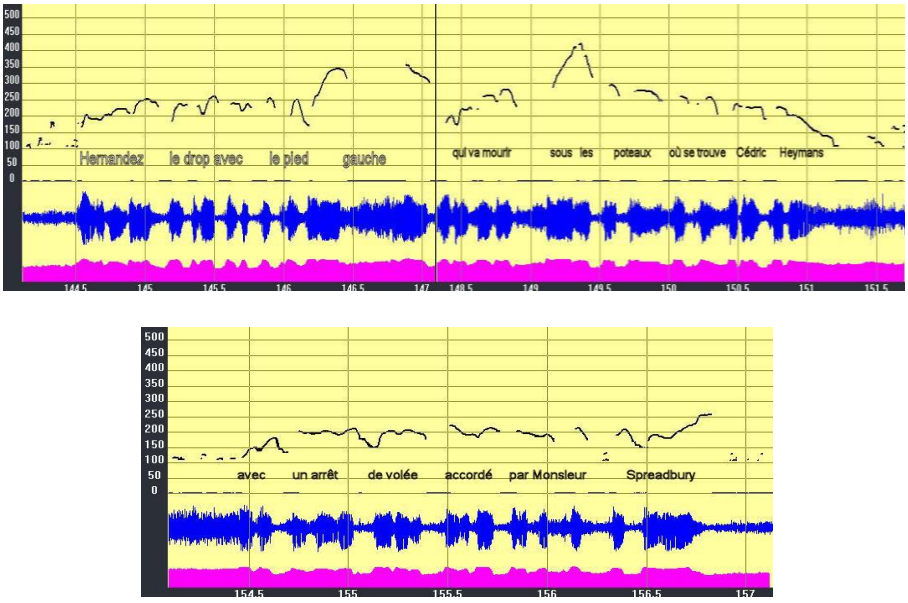


Figura 5: Illustrazione della frase: « Hernandez le drop avec le pied gauche qui va mourir sous les poteaux où se trouve Cédric Heymans avec un arrêt de volée accordé par Monsieur SpreadBury »

Ci occuperemo ora dell'inglese e del giapponese per mostrare in quale modo l'intonazione trasmette la stessa intenzione attitudinale nello stesso modo.

La stessa fase di gioco viene descritta in francese nella figura 6 e in inglese nella figura 7. Mentre tutto sembra "calmo" e i giornalisti "fanno passare il tempo" con un tono di voce relativamente basso (medio) e con poche variazioni, un giocatore francese commette un errore, perde la palla a vantaggio di un inglese che fugge e segna una meta: i giornalisti reagiscono nello stesso modo, passando ad un registro alto, accompagnato da un aumento della velocità di eloquio da parte del cronista inglese che manifesta così la sua eccitazione positiva, mentre il commentatore francese smozzica le parole come in una successione di singhiozzi.

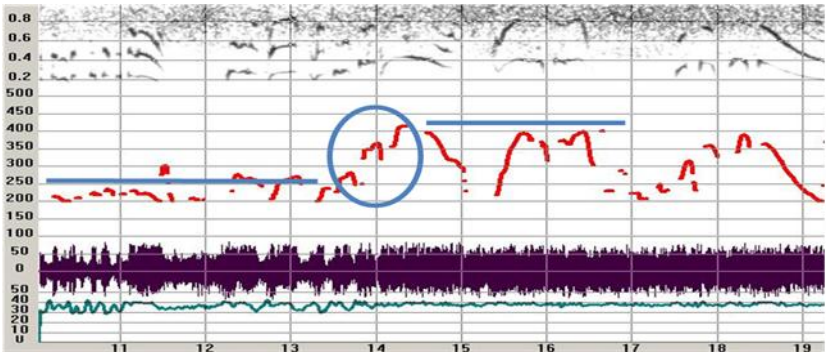


Figura 6: Gomarsall au pied c'est ce que vous expliquez tout à l'heure Thierry Lacroix la couverture (↗) de Traille (↘) # oh (↗) l'essai non # de Josh Lewsey (↗ ↘)…

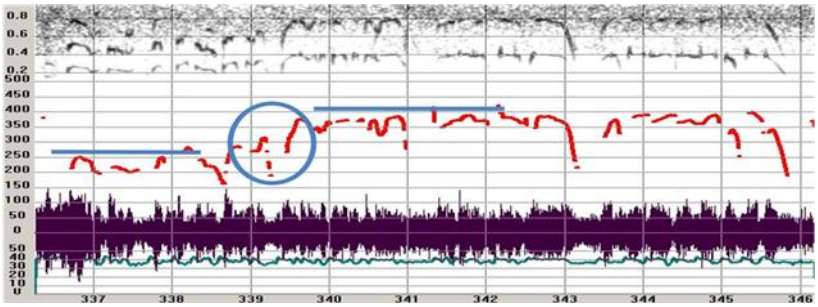


Figura 7: Traille has been in a quite central location of terrible (↗) hesitation from him…

Un fenomeno identico si riscontra per il commento di una stessa fase di gioco in giapponese e in francese.
Nell'ultimo minuto di gioco, quando il Giappone potrebbe segnare pareggiando, una successione di attacchi, segnalati da una successione di picchi della linea melodica, non ha esito e la formazione finisce col perdere la palla: questa azione mancata che segna la sconfitta del Giappone provoca una brusca caduta della F0 nel commento del giornalista, che segna al contempo la delusione e la fine dell'incontro.
La descrizione in giapponese è più rapida (il numero di more aumenta, segnato da una successione di brusche ascese della F0 prima della caduta finale (eccitazione "partigiana").

La descrizione in francese inizia con una velocità e una melodia "normali" (il giornalista è meno emotivamente coinvolto) e termina anch'essa con una brusca accelerazione associata a un'ascesa della F0 seguita da una caduta improvvisa e rapida su "fini".

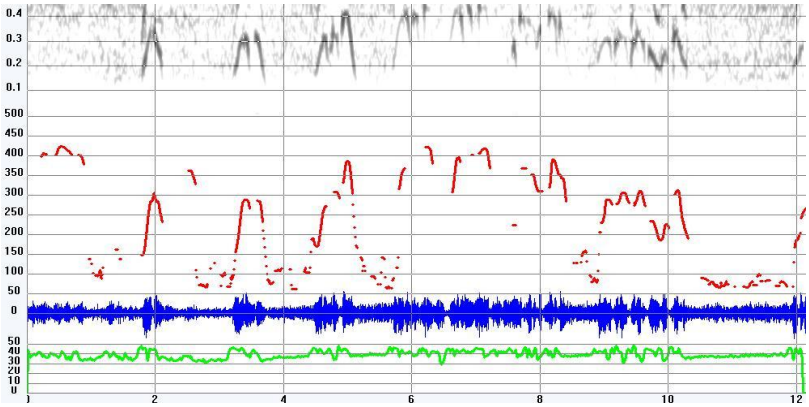


Figura 8: Robinsu ↗ OOnisi ↗ OOnisiga iku ↗ Oonisi itta usirokara hajikareta ↘ → aa korewa penaruti nokkuondesu ↘
(giocatore) (giocatore) (giocatore) (giocatore) (avanza)
(perde la palla) (ah, errore)

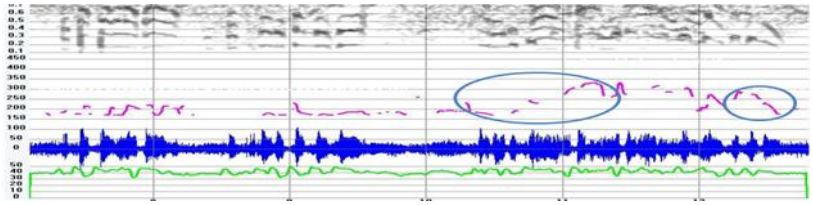


Figura 9: il faut pas perdre la balle # il faut créer des temps de jeu ## allez le soutien (↗) un ballon chipé par les Fidjiens c'est fini (↘)

5 Discussione

La nostra analisi dimostra che la cronaca del giornalista sportivo presenta due tipi di discorso: un discorso che è commentativo in senso proprio e un discorso descrittivo. Orbene: se dal punto di vista del contenuto del messaggio i due tipi di discorso sono facilmente identificabili (non foss'altro che a livello delle strutture sintattiche nominali che sono privilegiate nel discorso descrittivo), ciò è un po' meno

vero per quanto attiene alla prosodia. Infatti, già nel discorso commentativo si ritrovano degli aspetti che si sarebbero potuti considerare specifici del discorso descrittivo, in quanto riconducibili all'eccitazione motivata dall'azione di gioco in corso di svolgimento e descritta in diretta dal locutore. In questo senso, l'impiego di un registro di voce particolarmente alto, di accenti d'insistenza imprevedibili si presentano nei due tipi di discorso, descrittivo e commentativo, quando invece non erano previsti che nella parte descrittiva. Perché tale attesa? Poiché questo aumento del registro vocale, queste brusche variazioni della F0 sono associati, in linea generale, ad un discorso "emotivo". Si tratta infatti di parametri che si ripresentano spesso quando si descrivono emozioni come la gioia o la collera (Fonagy, 1983; Léon, 1993) che hanno in comune il fatto di essere emozioni attive e che sono associate all'idea di eccitazione. Evidentemente, l'esperienza ci insegna che queste particolarità della voce umana "eccitata" appaiono in un discorso realizzato in diretta, allorché il locutore è coinvolto dal gioco. Per contro, si fatica ad immaginare perché questi stessi tratti prosodici appaiano nel discorso prima che l'incontro abbia inizio.

È qui che bisogna di nuovo far riferimento alla situazione enunciativa per capire perché il locutore impiega un registro di voce piuttosto alto e tante variazioni della F0 nel suo discorso introduttivo, che è puramente commentativo. In effetti, bisogna ricordare che il cronista sportivo si rivolge prima di tutto al telespettatore, che non è in senso stretto presente nella situazione di enunciazione, ma che è costantemente tenuto in considerazione dal cronista nel suo discorso. La necessità di tener desta l'attenzione dei telespettatori giustifica l'impiego di un registro di voce alto, come pure la presenza ricorrente di accenti di insistenza volti a dare l'impressione che succeda qualcosa quando invece nulla è ancora cominciato. Non va quindi dimenticato che, anche se il locutore è coinvolto emotivamente nell'incontro, l'emozione è sempre, almeno in parte, una questione di messinscena volta a commuovere il telespettatore, a fargli condividere l'eccitazione del gioco e a tener desta la sua attenzione.

Dall'analisi delle realizzazioni in inglese e in giapponese ci si accorge che si ritrovano nelle tre lingue fenomeni prosodici analoghi e ciò indipendentemente dalla struttura fonologica e prosodica propria di ciascuna lingua. Il confronto fra diversi estratti di queste tre lingue è stato effettuato sulle parti puramente descrittive dei discorsi giornalistici. Infatti, abbiamo considerato che è nella parte descrittiva, quando il discorso è motivato dall'azione di gioco, che i corpora delle diverse lingue sono comparabili. Si può asserire che esiste una sorta di iconicità presente in questo discorso descrittivo nel senso che l'azione che si svolge contemporaneamente al discorso si riflette nella prosodia impiegata dal locutore. Abbiamo illustrato questo fenomeno nella figura 5. L'iconicità del discorso sportivo potrebbe probabilmente spiegare che, da una lingua all'altra, la descrizione di una medesima azione si traduca prosodicamente nello stesso modo.

Tuttavia, ciò non spiega per quale ragione questi stessi fenomeni prosodici appaiono quando il discorso cessa di essere descrittivo. Si potrebbe a questo punto ritenere di essere in presenza di un fonostile, cioè di una prosodia propria di questo tipo di discorso e che la mondializzazione degli scambi attraverso i media influenza anche il discorso giornalistico, unificandolo in un genere discorsivo e prosodico che condivide caratteristiche comuni in tutte le lingue, o perlomeno nelle tre lingue che abbiamo preso in considerazione. Questo fonostile potrebbe essere considerato come una sottocategoria del fonostile del discorso giornalistico con il quale ha in comune, perlomeno in francese, un accento all'inizio del gruppo ritmico (Lucci, 1980; 1983).

6 Conclusione

Il nostro studio si proponeva di offrire un'analisi e una descrizione dettagliata degli aspetti prosodici essenzialmente del discorso sportivo. Questa analisi è stata organizzata attorno a tre assi di riflessione: a) la specificità della situazione enunciativa, e segnatamente la presa in conto da parte del

locutore-giornalista nel suo discorso di un interlocutore invisibile e muto e tuttavia presente, il telespettatore; b) lo stretto legame esistente fra il discorso del giornalista (particolarmente nel discorso di tipo descrittivo) e l'azione di gioco; c) l'aspetto iconico di questo tipo di discorso che, a causa di ciò, presenta degli aspetti comuni in lingue molto diverse le une dalle altre a livello di struttura prosodica.

Abbiamo così dimostrato che, a causa della situazione di enunciazione e della necessità, per il locutore-giornalista, di mantenere viva l'attenzione del pubblico al quale si rivolge in realtà (pur non essendo questo stesso pubblico attore della situazione enunciativa), il registro vocale del locutore è praticamente mantenuto in modo permanente in un registro vocale piuttosto alto, con valori massimi della F0 assolutamente eccezionali, tenuto conto che siamo in presenza ogni volta di voci maschili. Inoltre, abbiamo mostrato che, per la stessa ragione, il locutore infiorava il suo discorso con accenti d'insistenza, imprevedibili, che si traducevano in bruschi salti melodici.

Analizzando il discorso descrittivo del locutore, abbiamo mostrato che esso presentava, dal profilo prosodico, dei legami con l'azione di gioco. In effetti, il discorso descrittivo del cronista presenta la stessa alternanza di fasi rapide e attive e di fasi più lente, il che corrisponde all'alternanza delle fasi di gioco durante l'incontro. Questo passaggio da una fase all'altra si traduce nel discorso del locutore attraverso variazioni di velocità e variazioni melodiche.

L'aspetto quasi simultaneo delle variazioni prosodiche in relazione con le variazioni del gioco ci hanno indotti a prendere in considerazione l'iconicità di questo tipo di discorso, tanto più che gli stessi fenomeni prosodici appaiono in tre lingue molto diverse quanto a struttura prosodica. La comparazione delle cronache sportive nelle tre lingue e le caratteristiche prosodiche comuni affiorate ci hanno indotti a riflettere sulla nozione di fonotipo applicabile in questo caso al discorso sportivo diffuso in tutto il mondo dai media.

Nel nostro studio, ci siamo interessati della cronaca di incontri di rugby. Pare quasi certo che ritroveremmo fenomeni

prosodici analoghi in altri commenti sportivi, perlomeno quando si tratta di sport collettivi. Resta il fatto che sarebbe certamente interessante il confronto con le cronache di incontri sportivi d'altro genere, come il calcio per esempio. Inoltre, abbiamo scelto di lavorare su cronache di incontri diffusi alla televisione; sarebbe senza dubbio interesssante studiare i vincoli legati al tipo di media attraverso il quale è trasmessa la cronaca. Incontreremmo forse strategie descrittive diverse da parte di giornalisti che si esprimono per radio e non più alla televisione, rivolgendosi a interlocutori dei quali sostituiscono anche la vista? Comunque sia, siamo qui in presenza di un genere di discorso, e al livello prosodico di un fonostile; un'analisi diacronica di cronache sportive radiofoniche e televisive potrebbe fornire un punto di vista interessante sulla creazione e sull'evoluzione di questo fonostile.

Ringraziamenti
Gli autori ringraziano Candido Matasci per la traduzione in italiano.

Riferimenti bibliografici

Blanche-Benveniste, C. (2000) *Approches de la langue parlée en français*. Paris: Ophrys.

Boulakia, G., J. Deulofeu and Ph. Martin (2001) Quand la prosodie bienforme les énoncés mal formés, in *Actes de la Journée Prosodie 2001 (CNRS-GDR 34 Phonologie),* Grenoble, 10-11 octobre 2001, pp. 53-6.

Deulofeu, J. (1998) Les commentaires sportifs constituent-ils un "genre", au sens linguistique du terme ?, in *Actes du Colloque Questions de méthode dans la linguistique sur corpus,* Perpignan, 9-11 Mai 1998, Paris: Champion.

Fónagy, I. (1983) *La vive voix*. Paris: Payot.

Kerbrat-Orecchioni, C. (1990) *Les interactions verbales*. Paris: Armand Colin.

Léon, P. R. (1993), *Précis de phonostylistique, parole et expressivité.* Paris: Nathan.

Lucci, V. (1980) L'accent didactique, in Fónagy, I. and P. Léon (eds) *L'accent en français contemporain*. *Studia Phonetica,* 15, pp. 107-22.

– (1983) *Étude phonétique du français contemporain à travers la variation situation nelle*. Grenoble: ULL.

Martin, P. (1981) Pour une théorie de l'intonation, in Rossi *et al* (eds) *L'intonation*, Paris: Klincksieck.

Griglie tonali

Amedeo De Dominicis
Università degli Studi della Tuscia – Viterbo (Italy)
dedomini@unitus.it

Abstract

The paper analyses some phonological processes of tonal and segmental perturbation, induced by the interaction among lexical tones (if there), syntactic/macrosyntactic architecture and intonational structures. Results bring evidence for a new intonation phrasing level (macroIP) and a new definition of intonation Nucleus. The data come from dialogical material, referring both to Italian (an intonational language) and Gizey (a tonal African language). The tonal and syntactic functions are represented by means of Tonal Grids.

1 Introduzione

Questo lavoro discute e propone un sistema di annotazione che intende rendere conto della dimensione sintattica e macrosintattica del parlato oltre che di quella tonale e intonativa. Il modello cui si ispira è quello delle griglie elaborate da Claire Blanche-Benveniste e dal GARS (*Groupe Aixois de Recherches en Syntaxe*). Ma, rispetto al suo antecedente, estende la gamma di fenomeni rappresentati.
Naturalmente, tale riferimento trova la sua giustificazione nella scelta del materiale analizzato e negli obiettivi metodologici del presente lavoro. Per quanto riguarda la prima questione, il sistema delle griglie nasce per dare adeguata rappresentazione alla produzione orale in una conversazione, dove la linearità è spesso interrotta, tanto che una data sequenza può essere realizzata per "onde successive". In particolare, serve bene ad evidenziare le disfluenze e il carattere frammentario del parlato (false partenze, esitazioni, ripetizioni) ed il loro contributo alla costruzione del significato e delle funzioni grammaticali.
Per quanto riguarda gli obiettivi, il presente lavoro propone un modello integrato sintattico-intonativo affinché, attraverso l'annotazione, emergano caratteristiche strutturali della dimensione intonativa finora poco studiate: in particolare, la ricorrenza (o *rima*) di particolari *patterns* intonativi caratterizzati da una doppia marcatura: perché ricorrenti e perché inattesi. Tali *patterns* sono, inoltre, collocati

su costituenti che possono essere ulteriormente determinati dal fatto che condividono identiche o simili funzioni sintattiche e/o alterazioni delle attese realizzazioni morfofonologiche a livello segmentale. Esploreremo, quindi, l'architettura del discorso alla luce di tali fatti segmentali e sovrasegmentali "grammaticalmente" inattesi e li considereremo come evidenze a supporto di una nuova visione del *phrasing* intonativo. In particolare, i costituenti che ospitano tali fenomeni imprevisti saranno considerati fortemente marcati nella suddivisione del testo in sintagmi intonativi (IP) e perciò saranno etichettati come Nuclei dell'IP, anche a prescindere dalla loro mera prominenza tonale. Queste unità saranno alquanto diverse dal classico Nucleo intonativo, anche se possono coincidere con esso, in quanto dominanti un sintagma più ampio del tradizionale IP, che chiameremo perciò *macroIP*.
I dati verranno estratti sia da una lingua intonazionale come l'italiano, che da una lingua tonale, come il gizey. Tale scelta intende sottolineare che il modello si applica ugualmente a tipi linguistici diversi. Infatti, i contorni intonativi e i toni lessicali sono entrambi espressioni della dinamica della F0. Differiscono solo per quanto riguarda l'estensione dei costituenti cui sono associati e per la gerarchia linguistica da cui dipendono. Nel primo caso, i contorni intonativi sono associati ai costituenti maggiori (come i sintagmi sintattici), mentre i toni lessicali sono una proprietà della sillaba; nel secondo caso, l'intonazione dipende dalla gerarchia fonologica, mentre i toni da quella lessicale (o da quella morfologica[1]). Nelle lingue tonali il tono lessicale e l'intonazione interagiscono, anche se si tratta di fenomeni allineati con unità segmentali diverse e la loro espressione superficiale è il risultato di due distinte gerarchie.

2 Il corpus

I dati in italiano sono presi dal corpus CLIPS[2]. Si tratta di un dialogo estratto dall'archivio relativo a Torino. In CLIPS è etichettato "DGmtA01T".

[1] In gizey i due toni manifestano l'opposizione tra aspetto verbale compiuto/incompiuto.
[2] Per gentile concessione del C.I.R.A.S.S., Università degli Studi di Napoli "Federico II". Il progetto CLIPS (*Corpora e Lessici di Italiano Parlato e Scritto*)

Il gizey è una lingua tonale con due toni fonologici (H, L), non è scritta ed è stata descritta parzialmente solo di recente (De Dominicis, 2007, in press; Melis and Ajello, 2008). È classificata nella famiglia ciadica, del gruppo afroasiatico (De Dominicis, 2002; 2003; Jungraithmayr, 1978; Pearce, 1998-99; Shryock, 1995; Wolff, 1983; 1987). In particolare, il gizey appartiene al ramo biu-mandara o masa, che include due sezioni: masa (masa, gizey) e musey (musey, ham, marba-lew, monogoy). Conta circa 12000 parlanti nativi nel Ciad meridionale e nel nord del Camerun. Nel luglio 2005 è stata registrata una conversazione tra due nativi. Parlano della costruzione della tipica casa di fango crudo, abitazione caratteristica delle popolazioni masa e gizey. Il luogo di provenienza di parlanti è lo stesso in cui la conversazione ha avuto luogo: Djougoumta, un villaggio rurale situato nella regione dell'Estremo Nord, dipartimento di Mayo-Danay, *arrondissement* di Guéré.

Sia i dati italiani sia quelli in gizey sono stati analizzati mediante il *software* di analisi acustica *Praat*. È stata estratta la F0 relativa alla traccia audio di ciascun parlante. Successivamente, i corpora sono stati annotati dal punto di vista sintattico. Per quanto riguarda l'annotazione intonativa dell'italiano, è stato adottato il sistema InTSInt (Hirst and Di Cristo, 1998; Hirst, Di Cristo and Espesser, 2000) nella versione automatizzata dello *script* denominato *mono_momel-intsint.praat*[3].

Mediante il sistema di annotazione in *tiers* e *TextGrid*, tipico di *Praat*, i due livelli di annotazione (sintattico e intonativo) sono stati allineati con il segnale audio. A partire da questa base di dati, sono state costruite le griglie tonali, illustrate di seguito. Invece, per il gizey sono annotati i toni lessicali, secondo l'indagine dettagliata di seguito (in 3.1.2.): i diacritici "↓" e "↑" descrivono le variazioni verso il basso e verso l'alto dei toni lessicali "H" e "L" (cfr. paragrafo 4)[4].

---

comprende la registrazione di materiale parlato e scritto della lingua italiana, realizzata in quindici aree geografiche diverse. Il parlato letto e semispontaneo è registrato tramite la tecnica del *Map-Task*; altre registrazioni comprendono parlato radiofonico, televisivo e telefonico.

[3] Copyright: Cyril Auran, 17 novembre 2004.

[4] I casi gizey sono stati già oggetto di una precedente pubblicazione (De Dominicis, in press). Per mancanza di spazio, non è possibile presentare le immagini relative ai contorni intonativi del corpus.

3 Metodologia di analisi qualitativa

La metodologia di analisi prende in considerazione l'insieme dei fatti segmentali e sovrasegmentali che non sono predicibili in base alle aspettative grammaticali di ciascuna delle due lingue esaminate. Cerca, inoltre, di fornire una spiegazione di tali fatti, alla luce di un'analisi all'interfaccia tra intonazione e sintassi. Si tratta, quindi, di un'analisi *dipendente* dalla lingua. Le evidenze considerate saranno, quindi, diverse per italiano e gizey. Di seguito, forniamo il dettaglio dei fenomeni considerati per ciascuna lingua.

3.1 Fenomeni del gizey

3.1.1 [ʔ] finale

Davanti a pausa, le radici nominali e verbali del gizey, che terminano in vocale, aggiungono un'occlusiva glottidale, oppure la vocale finale viene cricchiata o desonorizzata (De Dominicis, 2007). Il fenomeno è confermato da studi comparativi sui tipi di fonazione (Gordon and Ladefoged, 2001). Nel corpus talvolta questa marca fonotattica appare anche fuori contesto (cioè dopo una radice terminante in consonante) e talaltra non compare, invece, nei contesti in cui dovrebbe trovarsi (cioè dopo una radice terminante in vocale). In tali circostanze la marca o la sua assenza sono fatti inattesi. Perciò, essi rappresentano casi di studio che richiedono una spiegazione alla luce di un'analisi intonativa più complessa, in quanto la collocazione di tali fenomeni corrisponde ai costituenti sede di nucleo intonativo o di *perturbazioni tonali* nell'IP.

3.1.2 Perturbazione Tonale

In una lingua tonale i toni lessicali e il contorno intonativo interagiscono e il contorno realizzato è il risultato della F0 generata dai toni lessicali e dall'intonazione associata alla frase (cfr., per es., il modello addizionale di Gussenhoven, 2004). Perciò, i toni lessicali possono essere abbassati o innalzati rispetto a quanto atteso. Ci si riferisce a quei casi di *downstep* e di *upstep* che non sono motivati da cause fonologiche. Ad esempio, un *downstep* di tono alto preceduto da uno basso o un *upstep* di tono basso preceduto da uno alto vengono

considerati fonologicamente predicibili e, quindi, irrilevanti ai nostri fini. Mentre, quando questo innesco contestuale manca, allora la perturbazione del tono lessicale atteso è rilevante e può essere considerata una conseguenza del contorno intonativo. Perciò, è classificata come *perturbazione tonale*. Lo stesso accade quando i fenomeni di *downstep* e *upstep* in una data lingua risultano facoltativi. Questo è il caso del gizey, ove una sequenza di toni LH può anche non determinare *downstep* ed una HL può non innescare un *upstep* (cfr. caso 9). Per cui, classificheremo come perturbazioni tonali ogni occorrenza di *downstep* e *upstep*.
Occorre specificare che in questa ricerca il valore atteso di F0 di un dato tono lessicale è stabilito in riferimento al valore medio di quel tono pronunciato dal medesimo parlante in tre occorrenze della parola in cui si trova. Ogni parola che compare nel dialogo è stata fatta pronunciare dallo stesso parlante in una seduta di registrazione a parte. A questo scopo è stata usata una frase *carrier* nella quale la parola in questione compare in posizione interna. Ciò ha permesso di disporre di una base dati lessicali relativa allo *skeleton* tonale di ogni lemma usato nella conversazione che costituisce il nostro corpus.

3.1.3 Perturbazioni Segmentali

Oltre alle perturbazioni tonali, si possono osservare anche perturbazioni segmentali. A volte, i due tipi di perturbazione si collocano sui medesimi costituenti. Questi sono i casi cui ci interessiamo. Si tratta di cambiamenti di foni, per inserzione (per es. un [ʔ] finale di parola) o per sostituzione (per es. la realizzazione implosiva di una consonante esplosiva). L'interesse per tali fenomeni è motivato dal fatto che essi sarebbero inesplicabili se non ipotizzando un quadro di restrizioni che scaturiscono da un modello intonativo nel quale tali perturbazioni servono a marcare i costituenti sede di analoghe perturbazioni tonali.

3.1.4 Griglia Sintattica

La rappresentazione dei dati sintattici è condotta mediante *griglie*. Si tratta di una risorsa per descrivere le dipendenze sintattiche in un diagramma bidimensionale. È stata messa a punto dai lavori del *Groupe Aixois de Recherche en Syntaxe* (Blanche-Benveniste, 1979;

1990; 1997; Blanche-Benveniste *et al*, 1979; 1990; Bilger, 1982; Bilger *et al*, 1997). Come abbiamo già osservato nell'introduzione, la griglia rende conto dei fenomeni legati alle peculiari modalità di costruzione del significato e delle funzioni grammaticali, tipiche della produzione orale. È costituita di due dimensioni: l'asse orizzontale rappresenta la sequenza delle posizioni sintagmatiche o costituenti; mentre quello verticale mostra le eventuali diverse realizzazioni paradigmatiche che si trovano su una stessa posizione. La somma di una costruzione sintagmatica e dei suoi frammenti paradigmatici è detta *configurazione discorsiva*. Quest'ultima può ricorrere a intervalli regolari, come una rima, e dare così al discorso una sua specifica architettura.
In parallelo con tale approccio sintattico, collocheremo i dati intonativi nella griglia tonale. Si tratta del medesimo diagramma bidimensionale, chiamato però a rappresentare anche i tratti intonativi di un IP. Quest'ultimo può corrispondere a una Frase o a una Clausola. Similmente alla griglia sintattica, anche quella tonale serve ad evidenziare la ricorrenza dello stesso *pattern* tonale in diverse posizioni sintagmatiche o sull'insieme paradigmatico di costituenti appartenenti alla stessa posizione sintattica. In entrambi i casi, se una data perturbazione tonale ricorre, allora ogni sua istanza costituisce un'occorrenza di un *pattern* intonativo *in rima*. Queste intonazioni in rima possono agire sulla semantica della frase (caso 9), ma la loro funzione principale è assicurare la coesione conversazionale (caso 8). Talvolta, le unità in rima intonativa si realizzano su costituenti posizionati al di là dei confini di un IP (caso 8) o oltre i confini di turno di parola (casi 6, 10). In questi casi l'architettura intonativa della conversazione ed i suoi *pattern*s in rima forniscono supporto alla cooperazione dialogica tra gli interlocutori.

3.1.5 Interfaccia Sintassi-Intonazione

In alcuni casi (cfr. caso 7) certi *patterns* di perturbazione e rime intonative condivisi da serie di parole marcano l'appartenenza di tali insiemi paradigmatici alla medesima funzione sintattica, anche quando i membri di tali sequenze si trovano al di là dei confini dell'IP. In altre parole, in questi casi, la similitudine dei meccanismi perturbativi istituisce e garantisce l'interfaccia tra sintassi e intonazione: i

costituenti coinvolti marcano a livello superficiale l'identità soggiacente della loro funzione sintattica.

3.2 Fenomeni dell'italiano

3.2.1 Perturbazione Tonale

In una lingua intonazionale come l'italiano non esistono evidentemente toni lessicali. Perciò, la nozione di perturbazione tonale non può essere applicata a valori tonali grammaticalmente predefiniti. Pertanto, si considereranno come esempi di perturbazione tonale i casi in cui si rilevano rime intonative o incrementi di F0 particolarmente rilevanti su costituenti sintattici ripetuti o comunque sede di funzioni sintattiche condivise tra più costituenti.

3.2.2 Perturbazioni Segmentali

Le perturbazioni segmentali che saranno prese in considerazione per l'italiano riguarderanno casi di omissioni di parte della parola (casi 1, 4), esempi di Clausole dipendenti prive di complementatore (casi 1, 2, 3), frammenti discorsivi interpretabili come atipici per la collocazione dell'avverbio in posizione di modificatore di SN, o alternativamente per l'assenza di una Clausola dipendente relativa (caso 4). Infine, si considererà come perturbazione segmentale il caso di un articolo dapprima mancante e poi inserito per correzione dopo una preposizione (caso 5).

3.2.3 Interfaccia Sintassi-Intonazione e Griglia Sintattica

I fatti di mera prominenza intonativa sono sovracategorizzati interfacciandosi con eventi sintattici. A seguito di tale associazione, alcune prominenze vengono promosse nella gerarchia prosodica, in quanto dotate di duplice dominanza (intonativa e sintattica). A queste diamo il nome di *Nucleo* e denominiamo *macroIP* il costituente dominato. Il macroIP può corrispondere ad uno o più IP.
Un ruolo primario è assolto dalla ripetizione e correzione lessicale. In un caso (caso 1) essa serve a costruire il Nucleo. In un altro caso (caso 4) accade esattamente l'inverso: la ripetizione e correzione lessicale si accompagna con la degradazione gerarchica di un costituente

inizialmente prominente e che si ripresenta, invece, la seconda volta privo di prominenza. Nel caso 2 la ripetizione (senza correzione) con un progressivo incremento di F0 costruisce il Nucleo. Nel caso 3 la progressione della ripetizione di posizioni sintattico-lessicali e della loro prominenza intonativa porta due costituenti ad assurgere entrambi al rango di Nuclei. Infine, nel caso 5 si assiste alla costruzione di due Nuclei di cui uno, pur perdendo la sua prominenza intonativa nella progressione delle ripetizioni, conserva il suo ruolo dominante grazie alla sua interfaccia sintattica che lo consacra come Nucleo dislocandolo a destra della frase.

4 Casi di studio

I dati estratti dai corpora di italiano e gizey sono organizzati nelle griglie tonali come segue. Nei casi di italiano la prima riga trascrive ortograficamente il testo; nei casi di gizey essa rappresenta la trascrizione fonetica stretta. Le righe successive sono usate solo per il gizey. La seconda riga presenta la trascrizione larga (fonologica). La terza riga fornisce la traduzione parola per parola in italiano e la quarta la traduzione libera in italiano. Ciascuna griglia rappresenta una clausola (C) o un sintagma intonativo (IP).
L'annotazione tonale dell'italiano è realizzata mediante le etichette INTSINT[5]. Per il gizey si usano i simboli dei toni lessicali "H" e "L" con l'aggiunta di "↓" / "↑" che indicano *downstep* (o *downdrift*)[6]/*upstep* tonale. PROM indica un punto di prominenza tonale.

[5] Il sistema INTSINT codifica la curva intonativa come una sequenza di bersagli tonali (*targets*) la cui successione rappresenta il contorno di F0 in formato stilizzato. I simboli sono i seguenti: "H" (*Higher*) e "L" (*Lower*) per bersagli, rispettivamente, più alti e più bassi rispetto ai precedenti; "S" (*Same*) qualora il bersaglio sia allo stesso livello del precedente, e "D" (*Downstep*) e "U" (*Upstep*) per variazioni di minore entità, che spesso corrispondono a livellamenti della curva di F0 in tratti discendenti o ascendenti. Tre ulteriori etichette sono relative al *range* del parlante nell'unità tonale: "T" (*Top*) indica il valore più alto, "B" (*Bottom*) quello più basso, "M" (*Mid*) quello medio (tipico dell'inizio dell'unità intonativa).
[6] La distinzione *downdrift/downstep* si riferisce all'abbassamento del secondo di due toni H separati o meno da un tono L.

Per distinguere i turni di parola dei due interlocutori, si usano i simboli "P_1" e "P_2". I costituenti prominenti sono indicati con una sottolineatura della parola.

4.1 Caso 1

	sulla sua sinis…				
IP1	1	**P_1**	sulla	sua	sinis
			Prep	SN	
			M	PROM: S	B

	sulla tua sinistra				
IP2	1	**P_1**	sulla	tua	sinistra
			Prep	Agg	N
				PROM: T	B
				Nucleo	

In entrambi gli IP l'aggettivo possessivo occupa la posizione di massima prominenza intonativa (PROM). La funzione sintattica è la stessa, ma la copertura lessicale viene corretta ("sua" → "tua"). La prima PROM è espressa mediante il tono (M)**S**, la seconda mediante **T**. Quindi, alla correzione lessicale corrisponde un incremento di prominenza tonale della seconda PROM. Si tratta di un fenomeno che si trova frequentemente e che è notevole. La ripetizione lessicale serve da supporto alla costruzione di una PROM che si colloca sull'ultima occorrenza della serie ripetuta (e facoltativamente anche sulle occorrenze precedenti): nel corso della ripetizione la prominenza tonale si incrementa. Si tratta di un meccanismo di costruzione del Nucleo di natura linguistica particolare. Sebbene foneticamente entrambe le PROM siano prominenti, solo la seconda ("tua") è Nucleo intonativo nel quadro dell'ipotesi teorica che presentiamo in questo lavoro. E la natura di tale Nucleo è dovuta all'intersecarsi sul costituente "tua" di relazioni intonative e sintattiche. Per conseguenza, i due IP costituiranno un unico macroIP.

Caso 2

	quindi rivolta verso la torta						
IP1	1	**P_1**	quindi	rivolta	verso	la	torta
			Avv	V	Prep	Art	N
			M L	PROM: H L L	U	L	B

	sono sopra la torta					
IP2	1	**P_2**	sono	sopra	la	torta
			V	Prep	Art	N
			M	PROM: T	D	B

	no no devi stare sotto la torta							
IP3	1	**P_1**	no					
	2		no	devi	stare	sotto	la	torta
			NEG	V	V	Prep	SN	
			1: 2: M	U	PROM: H S	L	D D D D	

	a fianco alla torta					
IP4	1	**P_1**	a	fianco	alla	torta
			Prep	N	Prep	N
			M	PROM: H	L	L

	guardando la torta sulla sinistra					
IP5	1	**P_1**	guardando	la torta	sulla	sinistra
			V	SN	Prep	N
			M	PROM: U	L	S B
				Nucleo		

In questi cinque IP (e turni di parola) – pur se non consecutivi – il lemma "torta" ricorre in varie posizioni. Nei primi quattro non è PROM, mentre nell'ultimo IP lo diventa. Questa progressiva promozione nel rango della gerarchia prosodica è accompagnata e sorretta da una rima intonativa iniziale e da un successivo, graduale incremento di prominenza tonale. Infatti, nel primo IP di P_1 "torta" è associato al tono **B**, nel secondo (IP2 di P_2) il lemma viene ripreso – nonostante la distanza temporale che separa i due IP nella registrazione originale di CLIPS – nella medesima funzione sintattica (N in SPrep) e con lo stesso tono **B**. Dopo questa prima ripresa

lessicale con rima intonativa, "torta" ritorna nel successivo turno di parola di P_1, in cui è collocata sempre in posizione sintattica interna a SPrep, ma con tono leggermente più elevato: (L)**DDDD**. L'incremento di prominenza tonale prosegue, poi, nelle due successive riprese del lemma da parte di P_1: **L** e poi (M)**U**. Nell'ultima il lemma ha una posizione sintattica diversa (appartiene a SV) ma diventa PROM dell'IP, concludendo la sua ascesa nella gerarchia prosodica e raggiungendo il massimo valore di prominenza tonale.
Questa macrostruttura può essere interpretata funzionalmente come un meccanismo di coesione discorsiva intra- e inter-parlatore: la rima intonativa assicura la coesione tra i turni di parola di P_1; e quello di P_2, pur isolato da quelli di P_1, è nondimeno fortemente coeso, tanto da garantire l'architettura coerente del discorso di P_1.
Inoltre, a questa interpretazione funzionale si può aggiungere qualche considerazione relativa alla nozione di Nucleo nella teoria linguistica dell'intonazione. I dati mostrano che il Nucleo è costituito sulla base di un meccanismo di apparentamento tonale di natura linguistica e non solo attraverso una mera prominenza fisica. Nell'ipotesi teorica di questo lavoro, il Nucleo intonativo è solo il costituente "torta" in IP5. Su di esso si concentrano e si incrementano i fattori linguistici di prominenza tonale e di interfaccia sintattica. I cinque IP si riuniscono, quindi, in un unico macroIP governato dal Nucleo "torta".

Caso 3

	passando sopra gli sci ?				
IP1	1	**P_2**	passando	sopra	gli sci
			V	Prep	SN
			V	Modif.	Avv.
			PROM: M	D	B

	no gli sci io non ce li ho								
IP2	1	**P_1**	no	gli sci	io	non	ce	li	ho
			NEG	SN	PRO	NEG	Avv	PRO	V
			B	PROM: H	B	U		B	

	io passo sopra gli sci o no ?								
IP3	1	**P_2**	io	passo	sopra	gli	sci	o	no
			PRO	V	Prep	Art	N	Cong	NEG
			PROM: M	B	U	D	PROM: UU	B	
			Nucleo				Nucleo		

In questi tre IP (e turni di parola) si osservano due fenomeni paralleli di promozione di costituenti nella struttura gerarchica prosodica cui corrisponde un incremento di prominenza tonale. Il primo IP (di P_2) presenta il lemma "sci" in posizione sintattica di SPrep e con tono **B**. Nel secondo (di P_1) ha diversa funzione sintattica (interna di SV) ma con tono più alto (**H**) e funzione di PROM. Infine, nel terzo IP (di P_2) "sci" è di nuovo in posizione sintattica di SPrep (come nel primo IP) ed è anche PROM (come nel secondo IP) con tono **UU**. Parallelamente, un altro lemma segue un destino simile. Si tratta del pronome "io", che compare nel secondo IP (con tono **B**) ed è ripreso nel terzo IP (con tono **M**): nel secondo IP non è PROM, ma lo diventa nel terzo dove presenta un parallelo incremento di prominenza tonale. Quindi, i due lemmi ("sci", "io") seguono un destino parallelo: nel passaggio da un IP all'altro, da un interlocutore all'altro, incrementano sia il loro grado di prominenza tonale che la loro posizione gerarchica nella struttura prosodica. Nell'ultimo IP (IP3 di P_2) entrambi conseguono la posizione di PROM. Sicché, nello stesso IP coabitano due PROM.

Nell'ipotesi teorica del presente lavoro solo queste due ultime PROM ("io" e "sci" in IP3) rappresentano i veri Nuclei intonativi. Come nel caso precedente, la costruzione della loro posizione gerarchica è il risultato di un'architettura di rimandi sintattico-lessicali e di incrementi progressivi di prominenza, alla quale cooperano i due interlocutori. Ma il risultato in quest'ultimo caso è ovviamente più interessante e complesso, poiché lo sdoppiamento nucleare è fortemente marcato per quanto riguarda le predizioni della teoria del *phrasing* intonativo. I due Nuclei governano un unico macroIP costituito dai tre IP.

Caso 4

	la circonferenza prima				
IP1	1	**P_2**	la	circonferenza	prima
			Art	N	Avv
			H	PROM: T	B

	la semicirc… arrivava				
IP2	1	**P_2**	la	semicirc	arrivava
			Art	N	V
				PROM: H (L)	U D L

Si tratta di due IP consecutivi del medesimo locutore (P_2). Nel primo "circonferenza" è PROM con tono **T**. Nel secondo il lemma è corretto ("semicirc…") ed è sempre PROM, ma con tono meno alto (**H**). In breve, il parlante ripete e corregge. Il suo completamento è insieme lessicale e intonativo: alla maggiore specificazione lessicale si accompagna un decremento di prominenza tonale. In questo caso nessuna delle due PROM costituisce il Nucleo intonativo inteso nell'interpretazione che noi diamo a questa nozione. Perciò, il caso 4 è un esempio di un frammento di parlato in cui due costituenti, pur se fisicamente PROM, tuttavia non ricevono supporto linguistico per diventare Nuclei intonativi.

Caso 5

	fino a che metà del foglio					
IP1	1	**P_2**	fino a	che metà	del	foglio
			Prep	SN	Prep	N
			PROM: T	M	L	L S

	oltre la metà del foglio				
IP2	1	**P_2**	oltre	la metà	del foglio
			Prep	SN	SPrep
			PROM: U U	D	L

IP3	*eh no no arriva arriva <u>fino a</u> no alla <u>metà</u> del televisore*								
	1	**P_1**	eh	no					
	2			no	arriva				
	3				arriva	fino a	alla	metà	del televisore
	4			no					
			Interie.	NEG	V	Prep	Prep	N	SPrep
			M	1: 2: H 4: L	2: SU 3: DH	PROM: T	SU	PROM: T	DMB

IP4	*alla <u>metà</u> del televisore <u>fino a</u> qui ci sono*						
	1	**P_2**	alla metà del	televisore	fino a	qui	ci sono
			Prep	N	Prep	Avv	SV
			PROM: M	LB	PROM: H	B	
			Nucleo		Nucleo		

Nei quattro IP due lemmi vengono ripresi e trattati tonalmente dai due interlocutori in modo da giungere a costituire due Nuclei intonativi coesistenti nel medesimo IP e macroIP. L'operazione viene condotta cooperativamente e produce un effetto coesivo del discorso.
La preposizione "fino a" è la sede della PROM su tutti gli IP (e turni di parola), tranne il secondo. Nel primo IP il tono del Nucleo è **T**. Lo stesso tono viene ripreso in rima intonativa sul terzo IP, mentre sul quarto IP il tono diventa **H**. Parallelamente, il lemma "metà" è distribuito su tutti gli IP, ma diventa PROM solo negli ultimi due. Tale promozione nella gerarchia prosodica è espressa da una trasformazione dei toni associati alle occorrenze del lemma che termina con lo stesso tono con cui inizia: **M** → (U)**D** → **T** → **M**. I due **M** iniziale e finale esprimono il medesimo lemma, ma con posizione diversa nella gerarchia dei costituenti prosodici. Infatti all'inizio "metà" non è PROM, ma lo diventa alla fine.
Dunque, i due costituenti iniziano insieme il loro percorso strutturale nel primo IP (IP1 di P_2), dove occupano posizioni gerarchiche diverse (uno è PROM, l'altro no). Solo "metà" è ancora presente nel secondo IP (IP2 di P_2), con tono più basso e sempre in posizione non PROM. Interviene allora P_1 e nel terzo IP entrambi i lemmi sono ripresi e entrambi costituiscono PROM. Inoltre, entrambi sono espressi dal medesimo tono **T**. Tale rima intonativa istituisce una doppia PROM,

ma rinvia anche al tono **T** del precedente "fino" in PROM nel primo IP. Dunque, una duplice strategia di rima intonativa genera una gerarchia strutturale di costituenti prosodici molto marcata in quanto l'esito finale è un IP con doppia PROM. La spiegazione funzionale di tale doppia PROM appare chiara alla luce di quanto avviene nel successivo ed ultimo IP (IP4 di P_2). Qui il lemma "metà" è di nuovo PROM – così come lo è "fino a". Ma, mentre la posizione gerarchica preminente di "fino a" è supportata dalla sua prominenza tonale (tono **H**), quella di "metà" non si àncora ad una prominenza tonale (tono **M**): quindi, la doppia PROM nel terzo IP (di P_1) serve proprio a consentire la doppia PROM nel quarto ed ultimo IP (IP4 di P_2), sebbene in mancanza di una prominenza tonale di "metà". Ciò consente a P_2 di dislocare il costituente "metà" a destra rispetto a "fino a", diversamente da quanto avviene nei precedenti IP, e con l'obiettivo di restituire a "metà" quella salienza linguistica che ora, in termini di mera prominenza tonale, non possiede più.
In questo esempio, osserviamo un fenomeno che può essere spiegato supponendo una preprogrammazione tra i due interlocutori. Il destino strutturale di un costituente è guidato da entrambi i parlanti, i quali nei loro successivi turni di parola cospirano per spostare un costituente lungo la gerarchia strutturale. Il risultato è conseguito mediante la strategia linguistica della ripetizione lessicale, accompagnata da una cooperazione sul piano intonativo.
Per quanto riguarda la collocazione del Nucleo intonativo nella concezione dell'ipotesi su cu si basa il presente lavoro, i due costituenti ("metà" e "fino a") di IP4 sono i due Nuclei dell'intero frammento di parlato esaminato nel caso 5. Questi è un unico macroIP che raggruppa quattro IP.

Caso 6

Questo caso è il primo del corpus gizey. Insieme ad altri 6 frammenti della conversazione è stato oggetto di studio di un precedente lavoro (De Dominicis, in press).

C1 IP1	tʃ↓ápà k↓á? bl↓á tʃ↓á ŋòl↓ò? tʃápì ká báláwí ká ŋòl *aiuto* *esiste* *molto* *esiste* *molto* *c'era molto aiuto, molto*				
	1	**P_2**	tʃ↓ápà *aiuto*	k↓á? *esiste*	bl↓á *molto*
	2			tʃ↓á *esiste*	
	3	**P_1**			ŋòl↓ò? *molto*
			N	V	N
			SN	SV	
			↓HL	PROM: ↓H	↓H L↓L
			/i/ → [a]; V# !	1: [ʔ]# 2: /k/ → [tʃ]; V# !	1: V# ! [7] 3: [ʔ]#
				Nucleo	

In questo frammento due costituenti PROM costruiscono il Nucleo del macroIP, che coincide con i confini di IP. Ma l'intervento del secondo parlante (P_1) serve a rendere l'IP ben formato, chiudendolo con un abbassamento di valore della F0. Nei due turni di parola gli interlocutori (P_2 e P_1) cooperano al completamento della clausola (C) e dell'IP. SV è ripetuto: entrambe le occorrenze del verbo sono PROM e costruiscono insieme il Nucleo intonativo in rima. Il *downstep* (↓H) in rima è motivato fonologicamente dal precedente L o ↓H. Il completamento del parlante P_1 ricorre a un lemma diverso per esprimere "molto". Per spiegare la sua scelta, bisogna considerare il diverso *tonal skeleton* di [ŋòl↓ò?]: L↓L. I suoi toni L permettono la marcatura della fine dell'IP, perciò P_1 sceglie il lemma per chiudere l'IP. Una ripetizione di [báláwí] o [bl↓á], con toni H/↓H, non avrebbe consentito di ottenere lo scopo. Inoltre, sulla prima occorrenza del verbo "esiste" si trova una marca segmentale di fine parola tipica del gizey: [ʔ]#. Ma manca sulla seconda occorrenza (sebbene attesa) e si ritrova sulla seconda occorrenza di "molto" pronunciata da P_1 per completare la frase, mentre il primo "molto" (pronunciato da P_2) ne è privo, sebbene atteso. Insomma, i due costituenti PROM in rima, la scelta del materiale lessicale, la distribuzione dei *markers* fonotattici

[7] Il punto esclamativo, qui e di seguito, indica un fenomeno inatteso.

cospirano per lasciare aperta la struttura dell'IP, disponibile ad essere completata dall'intervento del secondo parlante: un esempio dell'esistenza di *patterns* cooperativi all'interfaccia tra sintassi e intonazione.

Caso 7

<table>
<tr><td rowspan="11">C1
IP1</td><td colspan="5">k↓áj n↓àm ɣ↓ájnʔ g↓ú péɗ↑àk↓íʔ kòjɗùmù k↑òjɗ↑ùmùʔ
kàjn nàm kàjn kú páɗàkí kòjdùmù kòjdùmù
DIM. esso DIMOST. REVERS. chiaro facile facile
questo è <u>facile</u></td></tr>
<tr><td>1</td><td rowspan="4">P₂</td><td>k↓áj
DIMOSTRATIVO</td><td></td><td></td><td></td></tr>
<tr><td>2</td><td>ɣ↓ájnʔ
DIMOSTRATIVO</td><td>n↓àm
esso</td><td>g↓ú
REVERS. [8]</td><td>péɗ↑àk↓íʔ
chiaro</td></tr>
<tr><td>3</td><td></td><td></td><td></td><td>kòjɗùmù
facile</td></tr>
<tr><td>4</td><td></td><td></td><td></td><td>k↑òjɗ↑ùmùʔ
facile</td></tr>
<tr><td></td><td rowspan="7"></td><td colspan="2">Modif.</td><td colspan="2">N</td><td>V (+SN)</td></tr>
<tr><td></td><td colspan="2">SN</td><td colspan="2">SV</td></tr>
<tr><td></td><td colspan="2">↓H</td><td>↓L</td><td>↓H</td><td>2: H↑L↓H
3: LLL
4: PROM:
↑L↑LL</td></tr>
<tr><td></td><td colspan="2">2: /k/ → [ɣ]; C[ʔ]#
!</td><td>C#</td><td>/k/ → [g];
V# !</td><td>2: /a/ → [e];
[ʔ]#
3: /d/ → [ɗ]; V#
!
4: /d/ → [ɗ];
[ʔ]#</td></tr>
<tr><td></td><td colspan="2"></td><td></td><td></td><td>4: Nucleo</td></tr>
</table>

In questo caso, IP e macroIP coincidono e, quindi, il Nucleo intonativo corrisponde alla nozione tradizionale. Esso si trova alla fine dell'IP, sulla seconda occorrenza dell'aggettivo "facile". Il SN ramificato da SV è costituito dalla ripetizione dell'aggettivo[9]. La prima ("chiaro") e la terza occorrenza ("facile") dell'aggettivo sono in

[8] Il morfema reversivo (o inversivo) indica una inversione completa di un'azione.
[9] La costruzione aggettivale equivale funzionalmente ad un SV.

rima intonativa: il loro *skeleton* di toni lessicali differisce nella prima e nella terza sillaba, ma nella seconda il medesimo tono L è realizzato come ↑L. In tal modo, l'unico tono lessicale che hanno in comune presenta la stessa perturbazione tonale. Essa è fonologicamente motivata come *upstep* innescato dal tono (H o ↑L) che precede ↑L. Inoltre, la seconda e la terza occorrenza dell'aggettivo mostrano la medesima perturbazione segmentale dell'occlusiva alveolare realizzata come implosiva; mentre la prima e l'ultima (la terza) occorrenza dell'aggettivo hanno in comune l'attesa marca di fine parola [ʔ]#. Questi due costituenti accerchiano la seconda occorrenza dell'aggettivo che, invece, è priva di tale marca fonotattica, sebbene del tutto attesa. Questa stessa marca di fine parola appare (ma stavolta inattesa) sul secondo *item* delle due occorrenze della particella dimostrativa all'inizio di IP: queste ultime due condividono anche il medesimo *pattern* di perturbazione tonale (↓H), che non è motivato fonologicamente, poiché tale *downstep* si realizza all'inizio di IP, dove al contrario ci si aspetterebbe un *reset* tonale innalzato. In tal modo, l'inizio e la fine di IP sono strettamente connessi: i costituenti all'inizio – così come quelli alla fine – di IP sono in rima intonativa reciproca, inoltre gli uni e gli altri sono reciprocamente relati per la distribuzione dei [ʔ]# attesi ed inattesi. In conclusione, due costituenti sintattici sono ripetuti (il dimostrativo e gli aggettivi); tale ripetizione segnala la loro appartenenza alla medesima funzione sintattica. Allo stesso modo, essi condividono gli stessi *pattern* di perturbazione tonale e le stesse marche fonotattiche attese/inattese. In altre parole, l'identità dei meccanismi perturbativi istituisce e si fa carico dell'interfaccia sintassi-intonazione: a livello superficiale, i costituenti interessati marcano l'identità soggiacente della loro funzione sintattica.

Caso 8

C1 IP1	kòʔ zìʔ zìʔ zìʔ zìʔ k *così casa casa casa casa DIMOSTRATIVO <u>costruire-essa</u> cioè così, la casa tu costruisci essa*						
	1	**P₁**	kòʔ *così*	zìʔ *casa*			
	2			zìʔ *casa*			
	3			zìʔ *casa*			
	4			zìʔ *casa*	kj↓énʔ *DIMOS.*	ínʔéʔ *costruire-essa*	àʔ *cioè*
			Avv.	N	Modif.	V	Avv.
				SN		SV	
			L	L	↓H	PROM: HH	L
			[ʔ]#	[ʔ]#	/a/ → [e]; C[ʔ]# !	/a/ → [ə]; [ʔ]#	[ʔ]#
						Nucleo	

C2 IP2	↓ín↓é ʔ↓áʔ ↓ínʔ t↓í ʔ↓è g↓àl vò vàj k↓éjn↓íʔ ínáʔ kàjnì ínáʔ tí gé gàl vò vàj kàjnì *<u>costruire-essa</u> DIM. costruire come lato in-casa nostra DIM. come <u>costruite essa</u> in nostro paese (gizey)*						
	1	**P₁**	↓ín↓é *costruire-essa*	ʔ↓áʔ *DIMOS.*			
	2		↓ínʔ *costruire*		t↓í ʔ↓è *come*	g↓àl vò vàj *lato in-casa nostra*	k↓éjn↓íʔ *DIMOS.*
			V	Modif.	Avv.	SPrep	Modif.
			SV		Avv.	SPrep	
			1: PROM: ↓H↓H 2: ↓H	↓H	↓H↓L	↓LLL	↓H↓H
			1: /a/ →[e]; V#! 2: C[ʔ]# !	[ʔ]#	V# !	V# !	/a/→[e]; [ʔ]#

	g↓àl gìzèjn k↓éjn↓íʔ ↓íné↓ʔ↓éʔ ↓ín t↓áʔ gàl gìzèj kàjnì ínáʔ ínáʔ tì gè *lato gizey DIM. <u>costruite-essa</u> costruite come ?* *dai gizey come <u>costruite</u> ?*					
C3 IP3	1	**P_1**	g↓àl gìzèjn *lato gizey*	k↓éjn↓íʔ *DIMOS.*	↓íné↓ʔ↓éʔ *costruire-essa*	
	2				↓ín *costruire*	t↓áʔ *come*
			SPrep		V-PRO	avv. interr.
			SPrep		SV	avv. interr.
			↓LLL	↓H↓H	1: PROM: ↓HH↓H 2: ↓H	↓H
			C#	/a/ → [e]; [ʔ]#	1: /a/ → [e]; [ʔ]# 2: C#	/e/ → [a]; [ʔ]#
					Nucleo	

Le tre occorrenze (nei tre IP) della particella dimostrativa condividono lo stesso *pattern* di perturbazione tonale (↓H). La medesima rima intonativa si presenta anche nella particella interrogativa finale ("come"). Inoltre, tutte le occorrenze del dimostrativo presentano la marca fonotattica di fine di parola, persino la prima, ove tale marca non dovrebbe apparire. In tal modo, la prima occorrenza rivela il suo apparentamento con le successive. Qualcosa di comparabile avviene a carico dei costituenti PROM (varie forme del verbo *costruire*): la marca fonotattica di fine parola ([ʔ]#) è presente nel primo e nel terzo PROM, scompare nel secondo, e ricompare – per quanto grammaticalmente inattesa – sulla sua ripetizione (nello stesso IP2). Questa mancanza opera come una freccia orientata verso il suo completamento, cioè verso i due altri PROM (quelli in IP2 e IP3 condividono la medesima perturbazione tonale).

Parallelamente, tutti i PROM e tutte le particelle dimostrative condividono anche alcune perturbazioni segmentali inattese. Tutti i costituenti associati alla posizione nucleare sostituiscono /a/ con [e] o [ə]. La stessa sostituzione segmentale interviene nei costituenti della particella dimostrativa che sono in rima intonativa; mentre, la particella interrogativa finale (che presenta una perturbazione tonale identica) inverte la sostituzione segmentale (/e/ → [a]). In breve, la sostituzione segmentale segue ed amplifica la relazione tra costituenti che sono già in reciproca relazione in quanto condividono un *pattern* intonativo comune. Nel complesso, tutti i fatti osservati cospirano

insieme al fine di rafforzare la coesione discorsiva delle tre Clausole ed IP che costituiscono il turno di parola.
Per quanto riguarda la collocazione del Nucleo intonativo nella concezione dell'ipotesi su cu si basa il presente lavoro, i due costituenti PROM ([ínʔəʔ] e [ˈ ínéʔˈéʔ]) di IP1 e IP3 sono i due Nuclei dell'intero frammento di parlato esaminato nel caso 8. Questi è un unico macroIP che raggruppa tre IP.

Caso 9

	l˥˩jn l˥˩jn h˥ár g˥àŋg˥ èʔ nèl tʃ˥òʔ ɲjn ɲjn hár gàngì nàl tʃòw *luogo luogo tondo giù così posizione* *in questa circonferenza qui*						
C1 IP1	1	**P₂**	l˥˩jn *luogo*				
	2		l˥˩jn *luogo*	h˥ár *tondo*	g˥àŋg˥ èʔ *giù*	nèl *così*	tʃ˥òʔ *posizione*
			N	Modif.	Prep	Avv.	N
			SN		SPrep		
			PROM: ↑L	PROM: ↑H	↑L↑L	L	↑L
			C#	C#	[ʔ]#	C#	[ʔ]#
			Nucleo				

Si tratta di un esempio di frase nominale: la testa nominale è ripetuta. Ed entrambi i costituenti ripetuti sono PROM, unitamente al modificatore dipendente (aggettivo) nello stesso SN. I tre costituenti PROM condividono lo stesso *pattern* di perturbazione tonale (↑), il quale è motivato dalla collocazione sintagmatica dei tre costituenti all'inizio dell'IP. Il *pattern* ripetuto di tale rima intonativa funge da supporto all'interfaccia con la semantica del discorso: la *somma* dei tre lessemi ("luogo", "luogo", "tondo") costruisce il significato "circonferenza", che non esiste nel lessico gizey.

Caso 10

<table>
<tr><td rowspan="9">C1
IP1</td><td colspan="9">gr↓óʔ nàl t↓áʔ k↓óʔ gr↓óʔ tʃ↓úl↓úm v↓áʔ d↓íʔ
gráwt nàltá kóʔ gráwt tʃúrúm fáɗ dì
mattino così là anche mattino <u>presto</u> immediatamente neg.
Anche al mattino, ma non immediatamente <u>presto</u> al mattino</td></tr>
<tr><td>1</td><td>P₂</td><td>gr↓óʔ
mattino</td><td>nàl
così</td><td>t↓áʔ
là</td><td>k↓óʔ
anche</td><td></td><td></td><td></td></tr>
<tr><td>2</td><td>P₁</td><td>gr↓óʔ
mattino</td><td></td><td></td><td></td><td>tʃ↓úl↓ú m
presto</td><td>v↓áʔ
immediatamente</td><td>d↓íʔ
neg.</td></tr>
<tr><td></td><td rowspan="5"></td><td>N</td><td>Avv.</td><td>Avv.</td><td>Avv.</td><td>Avv.</td><td>Avv.</td><td>NEG</td></tr>
<tr><td></td><td colspan="4">SN</td><td colspan="3">SN</td></tr>
<tr><td></td><td>↓H</td><td colspan="2">L↓H</td><td>↓H</td><td>PROM :
↓H↓H</td><td>↓H</td><td>↓
H</td></tr>
<tr><td></td><td>/a/ → [o]; [ʔ]#</td><td colspan="2">[ʔ]#</td><td>[ʔ]#</td><td>/r/ → [l]; C#</td><td>/f/ → [v]; [ʔ]#</td><td>[ʔ]#</td></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td>Nucleo</td><td></td><td></td></tr>
</table>

È una frase nominale, ma è divisa in due turni di parola. P_1 riprende il N di P_2 con lo stesso *pattern* intonativo (↓H). Inoltre, fornisce anche il costituente PROM che manca nel turno di parola di P_2. La stessa perturbazione tonale ricorre anche sulla particella negativa finale, che dovrebbe essere realizzata L, ma invece è ↓H. Infine, la marca fonotattica di fine parola ([ʔ]#) si trova su tutti i costituenti della frase (in entrambi i turni di parola). In breve, la distribuzione della marche fonotattiche lungo tutta la frase, i patterns intonativi in rima tra costituenti ripetuti, PROM e negazione finale, e – soprattutto – la mancanza di un costituente PROM nel primo turno di parola sono tutti indici la cui funzione è richiedere un completamento discorsivo. Il secondo turno di parola adempie a tale funzione di coesione discorsiva e di cooperazione conversazionale, fornendo il costituente PROM ed il *pattern* intonativo in rima. Alla luce di tali considerazioni, il costituente PROM è dotato delle funzioni linguistiche per assolvere il ruolo di Nucleo del MacroIP, costituito da due turni di parola.

5 Conclusioni

Il sistema di annotazione a griglia ha consentito di individuare coesioni linguistiche tra alcuni fenomeni collocati nella dimensione discorsiva e interpretati qui come emergenze superficiali di relazioni all'interfaccia tra i diversi componenti della grammatica: intonazione, sintassi, morfofonologia, lessico. I costituenti che ospitano questi intrecci di relazioni sono, quindi, le tessere disperse di un mosaico che trova la sua spiegazione nell'architettura profonda del discorso e nella coesione che tiene insieme i turni di parola del dialogo. L'analisi ha reso palese l'organizzazione soggiacente di tale rete di relazioni. Ha, inoltre, messo in evidenza che questi costituenti che emergono alla superficie del discorso carichi di complessi investimenti soggiacenti e di forte interdipendenza reciproca godono di proprietà atte a consentire loro di assolvere una funzione di testa nella segmentazione intonativa. Infatti, il peso delle relazioni paradigmatiche che in essi convergono e la forza della coesione sintagmatica espressa dalla loro ricorrenza nel testo rendono prominenti tali costituenti. Si tratta di una prominenza che scaturisce dalla densità delle funzioni linguistiche associate a tali costituenti. Quindi, può anche prescindere da una corrispettiva marca intonativa: questi costituenti possono non essere, perciò, segnalati da un movimento della curva intonativa di particolare ampiezza.

Nondimeno, la loro prominenza assegna a tali unità un ruolo nella gerarchia prosodica. Si tratta di capire quale. Nella classica rappresentazione del *phrasing* intonativo, ogni IP è dominato da un nucleo, individuato per la sua prominenza tonale. I costituenti qui definiti in base ad una prominenza di natura grammaticale saranno teste di sintagmi di estensione maggiore, che chiameremo macroIP.

Riferimenti bibliografici

Bilger, M. (1982) 'Contribution à l'analyse en grille'. *Recherches sur le français parlé* 4, pp. 195-215.

Bilger, M. *et al* (1997) 'La transcription de la langue parlée: quelques observations'. *Recherches sur le français parlé* 14, pp. 57-86.

Blanche-Benveniste, C. (1979) 'Des grilles pour le français parlé'. *Recherches sur le français parlé* 2, pp. 163-204.

– (1990) ‘Un modèle d’analyse syntaxique «en grille» pour les productions orales’. *Annuario de Psicologia* 47, pp. 11-28.
– (1997) *Approches de la langue parlée en français*. Paris: Ophrys.
Blanche-Benveniste, C. *et al* (1979) ‘Des grilles pour le français parlé’. *Recherches sur le français parlé* 2, pp. 163-205.
Blanche-Benveniste, C. *et al* (1990) *Le français parlé. Etudes grammaticales*. Paris: Editions du Centre National de la Recherche Scientifique.
De Dominicis, A. (2002) The Masa tonal system, in U. Gut and D. Gibbon (eds) *Typology of African Prosodic System. Proceedings*, Bielefeld: Universität, pp. 87-100.
– (2003) Tone or Pitch Accent in Masa?, in E. H. Wolff (ed.) *Topics in Chadic Linguistics. Papers from the 1st Biennial International Colloquium on the Chadic Language Family (Leipzig, July 5-8, 2001),*
Koeln: Ruediger Koeppe Verlag, pp. 53-72.
– (2007) [2006] Tonal patterns of Gizey, in A. De Dominicis (ed.) *Undescribed & Endangered Languages: the preservation of linguistic diversity. Proceedings of the International Conference*, Newcastle: Cambridge Scholars Press, pp. 60-163.
– (in press) ‘Tones and intonation in Gizey’. *Studi Magrebini* (Atti del *VIII Afro-Asiatic Congress*, Napoli 11-13 settembre 2008).
Gordon, M. and P. Ladefoged (2001) ‘Phonation types: a cross-linguistic overview’. *Journal of Phonetics* 29, pp. 383-406.
Gussenhoven, C. (2004) *The phonology of tone and intonation*. Cambridge: Cambridge University Press.
Hirst, D. and A. Di Cristo (eds) (1998) *Intonation Systems: A Survey of Twenty Languages*. Cambridge: Cambridge University Press.
Hirst, D., A. Di Cristo and R. Espesser (2000) Levels of representation and levels of analysis for intonation, in M. Horne (ed.) *Prosody: Theory and Experiment*, Dordrecht: Kluwer Academic Publishers, pp. 51-87.
Jungraithmayr, H. (1978) ‘The Zime dialect cluster (“Kado”, “Dari”) in Southern Chad: its verbal aspect system’. *Afrika und Übersee* 61, pp. 1-27.
Melis, A. and Ajello, R. (2008) *Dictionnaire Gizey-Français. Suivi d’une liste lexicale Français-Gizey*. Pisa: ETS.
Pearce, M. (1998-99) ‘Consonant and tone in Kera (Chadic)’. *Journal of West African Languages* 27, 1, pp. 33-70.

Shryock, A. (1995) ‘Investigating Laryngeal Contrasts: An Acoustic Study of the Consonants of Musey’. *UCLA Working Papers in Phonetics 89*, pp. 1-117.
Wolff, E. (1983) ‘Tonogenese in tschadischen Sprachen’. *Afrika und Übersee* LXVI, pp. 203-20.
– (1987) ‘Consonant-Tone interference in Chadic and its implications for a Theory of Tonogenesis in Afroasiatic’, in D. Barreteau (ed.) *Langues et cultures dans le bassin du lac Tchad*, Paris: ORSTOM, pp. 193-216.

Il parlato dei mass media: analisi multilingue del parlato dei telegiornali

Antonella Giannini e Massimo Pettorino
Università degli Studi di Napoli l'Orientale
agiannini@unior.it; mpettorino@unior.it

Abstract

The results of experimental surveys carried out on a corpus of spoken Italian on both radio and television show unequivocally that the rhythmic-prosodic characteristics of the speech used on radio and television news programmes have changed radically over the past fifty years. Several studies show certain marked changes that indicate a well-defined direction. It must be pointed out that the changes involved do not relate to the text of the utterance, but to the way it is delivered (articulation rate, speech rate, fluency, silence percentage and tonal range). Do the rhythmic/prosodic changes relate only to Italian, or, because of mass means of communication, are they true of other languages? Do different language systems share the same model of rhythmic/prosodic patterns? Can we be said to be approaching a global type of speech? The results of this experimental research carried out on television information speech in Italian, English, French, Slovenian and Japanese show a surprising parallelism in the development of these languages. In fact, the television news speech of the 1960's and up to the mid-1970's, compared with television news today, shows less speech rate, longer pauses, a lower level of fluency, an intonation trend which is less varied, and minimum differences in rate of articulation. In particular, the data reveal that recent television news, in all the languages analysed, are characterised by a drastic diminishing of the percentage of silence as compared to past ones. Fluency too, though in a different way, shows a net increase in today's television news. Another interesting fact common to all the television news broadcasts is the stability of the articulation rate. The fact is that the television news is read by professionals, the articulation positioning must be precise, the target audiences must be reached and the speaker must be clear and unambiguous.

1 Introduzione

Per parlato trasmesso si intende il parlato dei mezzi di comunicazione a distanza, radio, telefono, cinema, televisione, internet. Ovviamente la tipologia del parlato varia non solo tra i mezzi di comunicazione, ma anche all'interno di uno stesso mezzo. La radio e la televisione sono, tra i mezzi di comunicazione, quelli più eterogenei in quanto, anche all'interno di una stessa trasmissione si

può avere, ad esempio, un parlato letto o spontaneo, preparato o recitato, formale o informale, dialogico o monologico. Diverso è il caso dei radiogiornali e telegiornali in cui il parlato può essere considerato come una realizzazione intermedia tra parlato letto e parlato spontaneo in quanto il testo è scritto in funzione della produzione orale. Le caratteristiche di questo tipo di parlato, molte delle quali condivise dal parlato trasmesso in generale, sono, ad esempio, la distanza tra parlante e ascoltatore; la eterogeneità dell'ascoltatore, anonimo e indefinibile sia dal punto di vista sociale che culturale; la monodirezionalità del messaggio in quanto il destinatario non interagisce e riceve passivamente le informazioni; la scelta di forme linguistiche semplici, esplicite e dirette, allo scopo di far presa su un pubblico diastraticamente diversificato; l'uso di una lingua standard.

Grazie allo sviluppo tecnologico è possibile, oggi, fissare e conservare nel tempo, suoni e immagini, confrontarli diacronicamente tra loro e, dal confronto, è possibile formulare interessanti ipotesi non solo sull'evoluzione socio-culturale di un Paese, ma anche sui cambiamenti che avvengono nel modo di comunicare. La comunicazione è ormai diventata circolare, è un dare e ricevere senza sosta, senza limiti e barriere: popoli diversi possono, contemporaneamente, sentire e vedere le medesime cose trasmesse in Paesi lontani.

Il modo di comunicare è sicuramente cambiato nel tempo: una voce registrata negli anni '50-'60 è infatti facilmente riconoscibile come "antica", più difficile è definire le sue caratteristiche e individuare le differenze che la rendono così diversa dalle voci alle quali siamo oggi abituati. La maggiore difficoltà, quando si effettuano confronti di tipo diacronico, è nella scelta del materiale da analizzare. Se il materiale non è omogeneo, infatti, si rischia di attribuire al periodo di tempo trascorso qualcosa che invece è dovuto ad altre cause. Un corpus utile a tale tipo di indagine è quello dei radiogiornali e telegiornali in cui molti dei fattori che possono in qualche modo incidere sulle caratteristiche ritmico prosodiche dell'enunciato sono tenuti sotto controllo. Questo tipo di parlato trasmesso è infatti costituito da parlato letto, è rappresentativo della lingua standard, è altamente intelligibile, è orientato verso l'ascoltatore ed è prodotto da speaker professionisti nella medesima situazione contestuale.

2 Stato dell'arte

I risultati di alcune indagini sperimentali condotte su corpora di italiano parlato, sia radiofonico che televisivo[1] (Pettorino and Giannini, 1994, Giannini and Pettorino, 1999; 2005; Pettorino, 2002; Giannini, 2004; 2006) mostrano in maniera inequivocabile che le caratteristiche ritmico-prosodiche del parlato utilizzato nei radiogiornali e telegiornali sono cambiate profondamente negli ultimi cinquant'anni. La figura 1 rappresenta un modello di parlato di speaker radiofonici e televisivi (anni '50-'60 e '90).

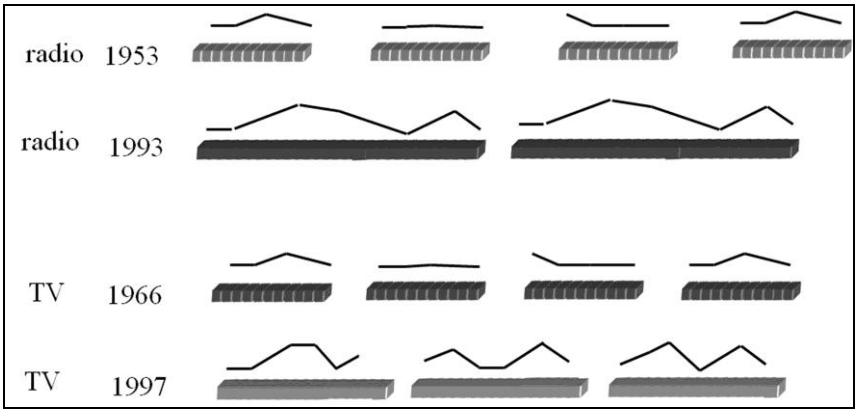


Figura 1: Schema ritmico del parlato degli speaker radio/televisivi degli anni '50-'60 e'90.

Come si vede, nel parlato di oggi le catene foniche[2] sono più lunghe, i silenzi sono meno frequenti e la loro durata diminuisce, il *range* tonale è più ampio ed è caratterizzato da un numero maggiore di picchi intonativi, mentre ciò che sembra non variare rispetto al passato è la durata sillabica. Vi è, quindi, una accelerazione della velocità di eloquio, quantificabile mediamente in una sillaba in più al secondo, che non incide sulla velocità di articolazione, il che significa che la

[1] Si tratta di lavori condotti con due diverse modalità: in alcuni lavori i brani di parlato letto avevano testi diversi, in altri, i testi messi a confronto, erano uguali.
[2] La catena fonica è la porzione di enunciato tra due silenzi successivi.

qualità articolatoria del parlato rimane accurata: l'accelerazione complessiva viene quindi ottenuta non a scapito della qualità ma piuttosto riducendo i silenzi. Tenuto conto delle differenze del mezzo, radio o televisione, si può dire che in entrambi i casi la tendenza sia univoca: si va verso un parlato più produttivo, nel senso che nello stesso tempo vengono pronunciate più sillabe. Questo risultato è ottenuto grazie a una diversa strategia nell'uso delle pause. Mentre una volta le pause erano nella quasi totalità sintattiche, e corrispondevano quindi a quello che nella scrittura è la punteggiatura, oggi le pause sono molto meno frequenti in quanto essenzialmente respiratorie: lo speaker utilizza pienamente la riserva d'aria inspirata senza tener conto del testo. Ma c'è di più: anche quando finalmente prende fiato, lo speaker di oggi impiega meno tempo grazie a una più attenta meccanica respiratoria. Il risultato è che nel parlato di oggi la componente "silenzio" si riduce spesso a meno del 10%.

Per quanto riguarda la durata delle pause silenti, gli enunciati prodotti nei vecchi telegiornali presentavano due tipologie diverse, una pausa lunga (circa 0,9 s) e una pausa breve (circa 0,4 s), rispettivamente corrispondenti ai confini sintattici forti (nella punteggiatura rappresentati dal punto) e deboli (rappresentati dalla virgola). Oggi si può dire che il punto sia scomparso, in quanto tutte le pause hanno grosso modo uguale durata (valore medio 0,3 s). Per quanto riguarda l'andamento intonativo si nota una maggiore varietà tonale, con un *range* molto più ampio. L'elemento che va sottolineato è l'utilizzo per così dire sintattico che viene fatto dell'intonazione. Poiché, come abbiamo visto, molte pause sintattiche sono scomparse, è necessario trasmettere l'informazione "confine di frase" in qualche altro modo. L'intonazione viene utilizzata a questo scopo. Nella maggior parte dei casi, infatti, immediatamente prima della pausa scomparsa compare un picco intonativo molto marcato, il cui scopo è evidentemente quello di far capire all'ascoltatore che in quel punto c'è una frattura, un confine. In questo modo la comprensione del messaggio è comunque salva. In più, si dispone di un tempo extra, quello della pausa soppressa.

3 Scopo del lavoro e corpus

I cambiamenti ritmico prosodici che abbiamo evidenziato riguardano soltanto l'italiano o, grazie ai mezzi di comunicazione di massa, lo stesso fenomeno investe altre lingue? Sistemi di lingue diversi hanno forse già trovato sul piano del ritmo e della prosodia un modello condiviso? Ed ancora, è possibile stabilire l'inizio di tali cambiamenti? Andiamo forse verso un tipo di parlato globalizzato?

Per rispondere a queste domande abbiamo proceduto ad un lavoro sperimentale in cui sono stati messi a confronto brani di parlato telegiornalistico di altri Paesi. Il procedimento utilizzato è analogo a quello seguito per l'italiano. Sono stati infatti selezionati brani tratti da telegiornali degli anni '60-'70 e confrontati con i corrispettivi telegiornali dei nostri giorni. Il corpus è costituito da brani in lingua slovena, inglese, francese del Canada e giapponese.

Le voci analizzate sono di speaker maschili tranne le voci del telegiornale sloveno. Per lo sloveno i brani esaminati, del 1977 e del 2009, sono tratti dalla rete nazionale RTV[3], per il francese l'emittente è la CBC[4] e i brani sono del 1970 e 2008, per l'inglese i brani, del 1963 e del 2008, sono della BBC[5] e infine per il giapponese l'emittente è la NHK[6] e i brani del 1978 e del 2008. Complessivamente sono stati analizzati circa 30 minuti di parlato. Per l'analisi è stato utilizzato il programma *Wavesurfer*. Per uniformare i risultati i dati ottenuti sono stati confrontati con quelli precedentemente ricavati per l'italiano (anni 1966 e 1997) in cui i testi dei telegiornali erano diversi.

Per ciascun telegiornale è stato computato il numero delle sillabe realmente prodotte e sono state misurate le durate delle singole catene foniche, le durate sillabiche, le durate delle pause silenti e il valore minimo e massimo di f_0 per ciascuna catena fonica. Successivamente, per ciascuna durata, sono stati calcolati i valori medi e ricavati i seguenti indici prosodici: velocità di articolazione (**VDA**), velocità di eloquio (**VDE**), fluenza (**F**), percentuale di silenzio (**PS**) e range tonale (**RT).**

[3] Radio Televizija Slovenija
[4] Canadian Broadcasting Corporation
[5] British Broadcasting Corporation
[6] Nihon Hōsō Kyōkai

La **VDA** equivale al rapporto tra il numero di sillabe e il tempo impiegato a produrle (sill/s). Essa non tiene conto dei tempi di pausa e il suo variare dà la misura dell'accuratezza del gesto articolatorio.

La **VDE** equivale al rapporto tra il numero di sillabe e il tempo complessivo dell'enunciato (sill/s) e pertanto tiene conto delle pause silenti. Il suo variare indica la maggiore o minore produttività dell'eloquio.

F corrisponde al numero di sillabe prodotte tra un silenzio e il successivo. **PS** è il valore percentuale di silenzio sull'intero enunciato e **RT** corrisponde all'intervallo tra valore massimo e minimo della frequenza fondamentale.

4 Risultati e discussione

Nella tabella I sono riportati i valori medi delle durate sillabiche, delle catene foniche e dei silenzi per ciascuna lingua analizzata e della lingua italiana di confronto.

I dati mostrano che la durata media delle catene foniche, negli telegiornali odierni, aumenta in maniera rilevante per l'inglese (da 1,74 a 3,3 secondi) e per il giapponese (da 1,2 a 1,95 secondi). L'incremento rilevato per l'inglese è simile a quello dell'italiano (da 1,76 a 3,80 secondi). Per quanto riguarda il giapponese, anche se il valore del 2008 è il più basso in assoluto, l'incremento rispetto al 1978 è di circa il 60%. In sloveno e francese non vi sono differenze significative (in entrambe le lingue la durata media delle catene foniche, per i telegiornali antichi e moderni, è di circa 3 secondi. Se consideriamo che i telegiornali antichi dell'italiano e dell'inglese risalgono agli anni '60, e quelli dello sloveno e del francese agli anni '70, si può avanzare l'ipotesi che, relativamente a questo parametro, gli anni '70 segnano l'inizio dei cambiamenti ritmico-prosodici nella comunicazione del parlato telegiornalistico occidentale.

Anche la durata media dei silenzi sembra mostrare uno stesso andamento, unica eccezione il giapponese, dove la pausa silente aumenta, anche se di poco, rispetto al 1978. Per le altre lingue il confronto, sia all'interno della stessa lingua sia con l'italiano, mostra negli odierni telegiornali non solo una riduzione in durata intorno al 40%, ma anche un valore medio comune di circa 0,3 secondi. Tale dato, come è stato rilevato nei lavori già citati, indica che, contrariamente a quanto avveniva negli enunciati prodotti nei vecchi

telegiornali in cui erano presenti due tipologie di pause (lunga e breve), negli odierni telegiornali le pause hanno tutte una durata breve.

Voci femminili	Sloveno 1977	Sloveno 2009
Dur. sillabe	0,166	0,178
Dur. cat. fon.	3,3	3
Dur. silenzi	0,577	0,344

Voci maschili	Inglese 1963	Inglese 2008
Dur. sillabe	0,222	0,210
Dur. cat. fon.	1,74	3,3
Dur. silenzi	0,554	0,331

Voci maschili	Francese 1970	Francese 2008
Dur. sillabe	0,192	0,161
Dur. cat. fon.	2,71	2,61
Dur. silenzi	0,405	0,300

Voci maschili	Giapponese 1978	Giapponese 2008
Dur. sillabe	0,125	0,126
Dur. cat. fon.	1,218	1,95
Dur. silenzi	0,460	0,508

Voci maschili	Italiano 1966	Italiano 1997
Dur. sillabe	0,150	0,170
Dur. cat. fon.	1,760	3,80
Dur. silenzi	0,400	0,290

Tabella II: Media delle durate

In figura 2 vengono riportate le percentuali di silenzio nei telegiornali antichi e moderni. Come si vede i telegiornali recenti, in tutte le lingue analizzate, sono caratterizzati da una drastica diminuzione della percentuale di silenzio rispetto a quelli più antichi. Si tratta, quindi, di un parlato molto più produttivo, in quanto il tempo a disposizione viene utilizzato più efficacemente. La maggiore differenza si ha nelle notizie dei telegiornali inglesi in cui il tempo di pausa passa dal 24% nel 1963 al 5,4% nel 2008. È interessante notare che la tendenza al modello condiviso non azzera, comunque, le differenze socio-linguistiche e socio-culturali: la maggiore durata delle pause nel giapponese ne è un esempio. Come dice Herbig[7], i giapponesi, al contrario degli americani, parlano poco e fanno pause più lunghe, usando un registro altamente formale. Le pause più lunghe degli speaker televisivi rilevate dal confronto confermano tale affermazione.

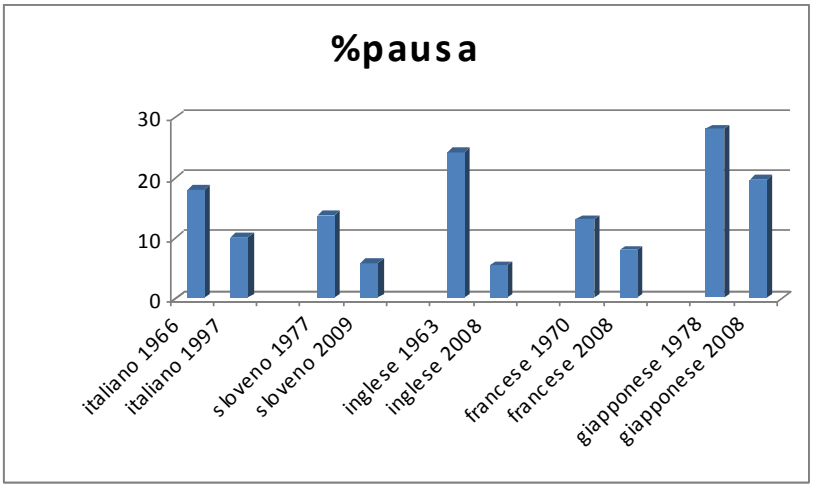


Figura 2: Percentuale di silenzio nei telegiornali "antichi" e "moderni"

La figura 3 mette a confronto i valori relativi alla VDA[8] per tutti i telegiornali esaminati. Come si vede i valori di VDA variano in relazione al sistema della lingua: i valori più bassi intorno alle 4,5

[7] Herbig 2003

[8] Quando la **VDA** aumenta, la durata media sillabica diminuisce, con conseguente riduzione del gesto articolatorio, causando un parlato affrettato e poco accurato, in cui spesso il target articolatorio non viene pienamente raggiunto.

sillabe/s sono quelli relativi all'inglese, i valori più alti, fino a 8 sillabe/s sono quelli relativi al giapponese. Tali oscillazioni dipendono dalla complessità della struttura sillabica, dovuta alla maggiore frequenza di occorrenza di sillabe aperte in giapponese e in italiano e di sillabe chiuse in inglese, sloveno e francese. Per quanto riguarda il confronto diacronico all'interno di ciascuna lingua, si notano delle variazioni non significative (non rilevanti sul piano percettivo) l'unico caso in cui vi è una accelerazione articolatoria rilevante è il caso del francese in cui vi è un incremento di 1 sillaba al secondo. Va detto che la relativa stabilità della VDA riflette la particolare tipologia di parlato, vale a dire un parlato letto prodotto da speaker professionisti. In altre parole un parlato trasmesso, accurato nel gesto articolatorio in quanto vengono raggiunti i target articolatori per ciascun segmento.

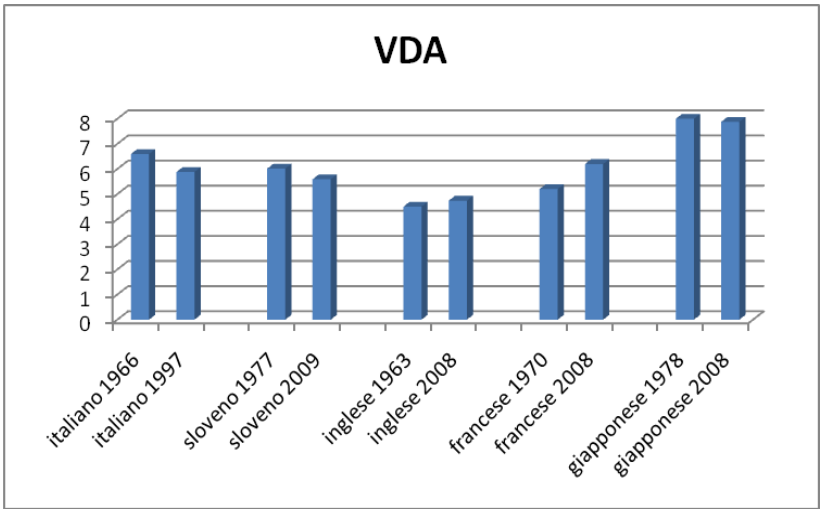


Figura 3: Confronto tra le velocità di articolazione

Anche la fluenza (figura 4), seppur in maniera diversa, presenta negli odierni telegiornali un incremento. Gli indici di fluenza che maggiormente si avvicinano a quelli rilevati per l'italiano (circa 6 sillabe), si riferiscono all'inglese (circa 8 sillabe) e al giapponese (circa 6 sillabe). L'incremento di circa due sillabe si ha nello sloveno e nel francese. Ancora una volta, in base a quanto ipotizzato in precedenza per la durata media delle catene foniche, agli anni '70 si può far risalire l'inizio dei cambiamenti dello stile del parlato dell'informazione. Conseguenza della più alta fluenza è un parlato più

produttivo: il tempo a disposizione è ridotto e va quindi sfruttato nella maniera migliore.

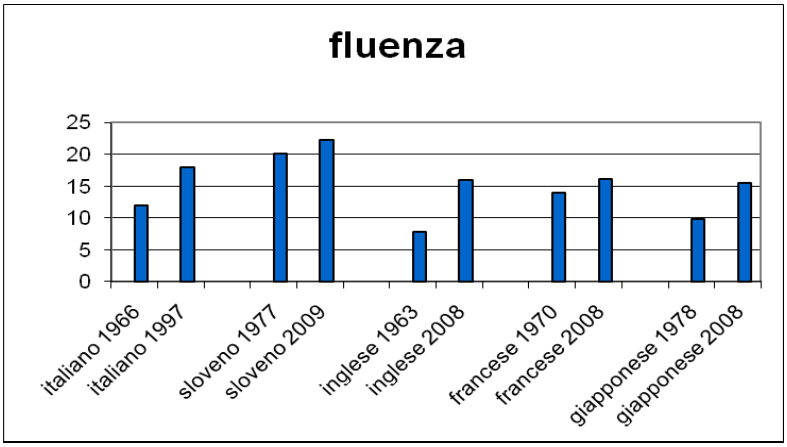


Figura 4: Indice di fluenza

Per quanto riguarda la **VDE,** le cui variazioni avvengono soprattutto in funzione del tempo di pausa (più frequenti e più lunghi sono i silenzi, più numerose sono le esitazioni, più basso è il suo valore), si nota uno stesso andamento (figura 5). Infatti in tutte le lingue esaminate la VDE aumenta in maniera costante, anche se in misura più o meno accentuata.

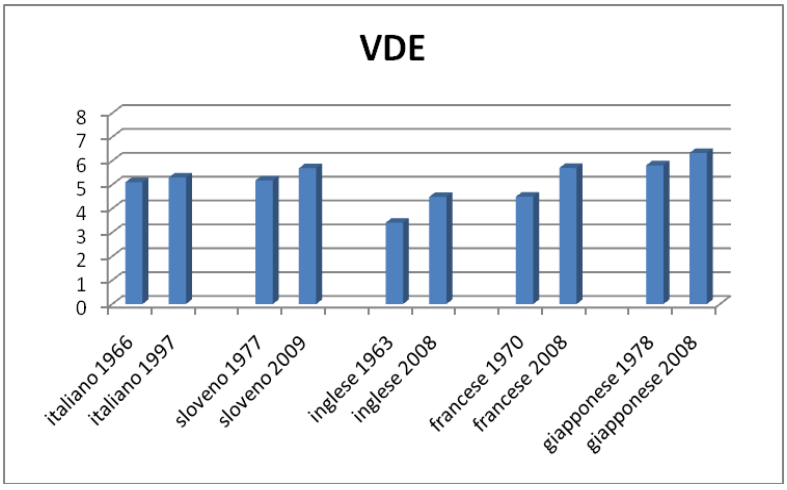


Figura 5: Velocità di eloquio

In figura 6 si può notare come, indipendentemente dalla frequenza fondamentale propria di ciascun parlante, gli odierni telegiornali siano caratterizzati da un *range* tonale più ampio. Il valore massimo di f_0 aumenta per ciascun *range* tonale, grazie all'uso sintattico dell'intonazione e dei ricorrenti picchi intonativi. Particolarmente evidente è il cambiamento diacronico per lo sloveno, dove il *range* aumenta di oltre un'ottava.

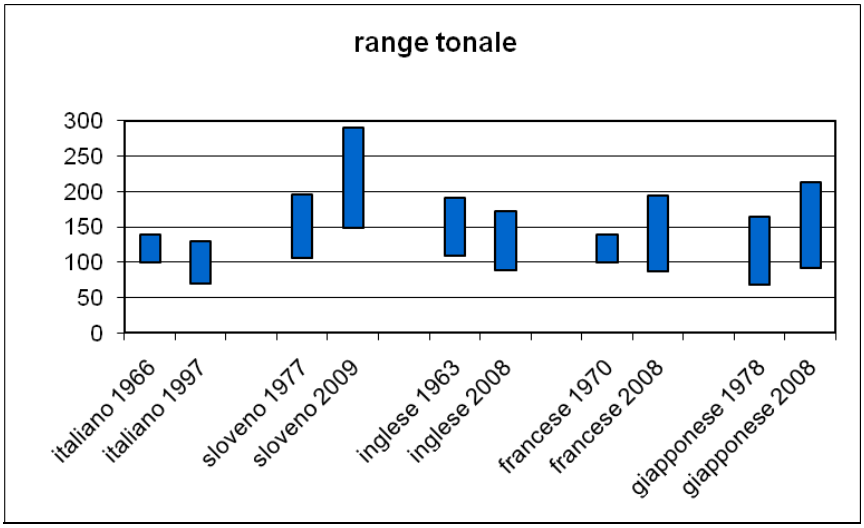


Figura 6: *Range* tonale

5 Conclusioni

I risultati di questa ricerca sperimentale condotta sul parlato dell'informazione televisiva per l'inglese, il francese del Canada, lo sloveno e il giapponese, confrontati con l'italiano, mostrano un sorprendente parallelismo dell'evoluzione di tali lingue seppure con alcune eccezioni. Il confronto diacronico ha infatti evidenziato che i moderni telegiornali, in tutte le lingue esaminate, sono caratterizzati da:

- catene foniche più lunghe
- un più alto indice di fluenza
- una maggiore velocità di eloquio
- pause più brevi

- una minore percentuale di silenzio
- una pressoché stabile velocità di articolazione
- un più ampio *range* tonale

L'ipotesi di un modello ritmico prosodico intonativo condiviso dal sistema globalizzato dell'informazione televisiva viene dunque confermato dall'analisi sperimentale. Sarebbe opportuno verificare, mediante ulteriori indagini, se tale modello sia limitato esclusivamente al mondo dell'informazione oppure stia assestandosi anche in altri settori della comunicazione.

Riferimenti bibliografici

Giannini A. (2004) Analisi acustica del parlato televisivo, in P. Cosi (ed.), *Misura dei Parametri, aspetti tecnologici ed implicazioni nei modelli linguistici*, *I convegno Nazionale dell'Associazione Italiana di Scienze della Voce (AISV*), Padova: EDK, pp. 49-61.

– (2006) I silenzi del telegiornale, in M. Pettorino, A. Giannini, M. Vallone and R. Savy (eds) *La Comunicazione Parlata*, Atti del Congresso internazionale (Napoli 23-25 febbraio 2006). Tomo I, eBook. Napoli: Liguori editore.

Giannini, A. and M. Pettorino (1999) I cambiamenti dell'italiano radiofonico negli ultimi 50 anni: aspetti ritmico-prosodici e segmentali, in R. Delmonte and A. Bristot (eds) *"Aspetti computazionali in fonetica, linguistica e didattica delle lingue: modelli e algoritmi", Atti delle 9^{e} Giornate di Studio del GFS*, vol. XXVI, Venezia, pp. 65-81.

Herbig P. A. (2003) *Marketing interculturale*, Apogeo, Milano,

Pettorino, M. (2002) I cambiamenti della lingua italiana, in A. De Dominicis (ed.) *La voce come bene culturale*, Roma: Carocci, pp. 141-157.

Pettorino, M. and A. Giannini (1994) Aspetti prosodici del parlato radiofonico, in P.L. Salza (ed.) *"Gli aspetti prosodici dell'italiano", Atti delle 4^{e} Giornate di Studio del GFS*, Torino, vol. XXI, pp. 19-28.

– (2005/2006) Italian TV speech: a diachronic analysis, *IX Simposio International de comunicación social, Santiago de Cuba 24-28 de*

enero de 2005, actas-II, Santiago de Cuba: Centro de Lingüística Aplicada, pp. 849-853 [ora in E. Miyares Bermúdez and L. Ruiz Miyares (eds) Linguistics in the Twenty First Century, Newcastle: Cambridge Scholars Press, pp. 103-111].

Il ruolo dell'intonazione nell'interpretazione dell'enunciato: informazioni categoriali o categorie informative?

Barbara Gili Fivela
Università del Salento, Lecce
barbara.gili@unisalento.it

Abstract

The paper reports results of production and perception experiments performed on the variety of Italian spoken in Pisa, in order to shed some light on the role of intonation in the linguistic interpretation of the message.

The analysis of read and semi-spontaneous speech data show that speakers usually have more than one choice of pattern to implement for a given function and meaning, besides the quite wide range of variation in implementation within the same pattern. These aspects are not interpreted as neglecting a phonological role of intonation. Rather, the lack of a one-to-one relation between intonation pattern and function and the observed variability are considered here to relate to categories in which members differ for expressing shades of meanings together with other linguistic information. Indeed speech perception data - categorical perception, imitation and card game tasks - show that meanings appear to be clearly identified at the extremes of gradient variations, but not to be differentiated in a clear-cut and abrupt way, as would be expected in the case of linguistic, categorical information. That is, tonal events distinguish meanings and functions but they do not necessarily correspond to discrete units. It is then argued that all the tonal events that participate in differentiating functions belong to categories that are part of the linguistic system and convey many different types of information. They are simply used differently from other linguistic (phonological) categories to exploit the related communication channel more effectively.

1.1 Introduzione

Le categorie linguistiche sono composte da elementi che si comportano in modo analogo o che svolgono una funzione simile ad un livello specifico della grammatica. Nelle trattazioni linguistiche si individuano categorie sintattiche, lessicali, morfologiche e semantiche (ad esempio, sintagmi, parti del discorso, genere, numero, parole interrelate). Queste categorie possono essere viste come insiemi omogenei e discreti oppure come insiemi che includono elementi più o meno rappresentativi della categoria stessa - si veda la teoria dei prototipi (Rosch 1973; 1975). A livello fonologico, le categorie sono normalmente intese come discrete e

aventi caratteristiche fonetiche 'stabili'. Infatti, in base alla 'quantal theory of speech' (Stevens 1972; 1989; per un contributo recente, si veda Stevens e Keyser 2010), le categorie possono essere viste come 'quantal regions', nettamente distinte le une dalle altre e caratterizzate al loro interno da parametri acustici e uditivi relativamente stabili, anche rispetto a cambiamenti nei parametri articolatori. Ne consegue che all'interno della categoria un cambiamento di tipo articolatorio non avrà un effetto apprezzabile, mentre lo avrà nella regione 'non stabile', producendo un drastico cambiamento nell'output acustico (che corrisponde, infatti, a una percezione categorica). In questo senso la relazione tra parametri articolatori ed output acustico (o effetto uditivo) non è lineare.

In letteratura la differenziazione acustico-articolatoria di vocali e consonanti è anche messa in relazione alla presenza/assenza di tratti distintivi e la discussione relativamente al loro essere caratterizzati in termini di 'quantal theory' è oggetto di indagine (ad es., Clements e Ridouane 2006). La definizione fonetica dei tratti distintivi delle categorie, infatti, non è ancora chiara e gli aspetti percettivi forniscono un quadro composito. Infatti, mentre la percezione delle consonanti sembra avvenire in modo categorico, in base alla minima variazione acustica nella regione 'non stabile' prevista al confine tra categorie diverse (Lieberman et al. 1957, sulle occlusive /b/, /d/, /g/), la percezione delle vocali sembra caratterizzata da soglie meno definite. Nel commentare i risultati per le vocali /ɪ/, /ɛ/, /æ/, Fry et al. (1962) fanno riferimento al fatto che, rispetto a quanto accade per le consonanti, manca per le vocali una vera e propria discontinuità articolatoria: la possibilità di categorizzazione percettiva potrebbe quindi dipendere dal fatto che i foni siano prodotti categoricamente (con discontinuità articolatoria, come le consonanti) o non categoricamente (ossia senza discontinuità articolatoria, come le vocali).

L'intonazione rappresenta un ambito di particolare interesse in relazione a questo argomento. L'analisi dell'intonazione è approfondita da tempo sia su un piano fonetico che su un piano fonologico, con il riconoscimento di eventi tonali che svolgono chiare funzioni linguistiche e presentano caratteristiche distinte da un punto di vista fonetico, senza che però si possa parlare di discontinuità articolatoria. Ad esempio, il lavoro di Bruce (1977) ha mostrato che la differenza di allineamento di eventi tonali determina una diversa interpretazione in relazione a due accenti lessicali in svedese. A partire da questa osservazione è stato proposto che anche per una lingua come l'inglese, nella quale le modificazioni di tono non

sono legate a differenze nel significato lessicale, fosse possibile individuare eventi tonali dalle caratteristiche distinte, facenti parte di un sistema fonologico (Pierrehumbert 1980 e l'intero quadro teorico della teoria autosegmentale-metrica, alla quale si fa riferimento in questo lavoro). Tuttavia non tutti gli studiosi ritengono che esistano categorie intonative di tipo fonologico in lingue come l'inglese, o come l'italiano, in cui l'intonazione opera a livello post-lessicale (Marotta, 2003). Infatti il problema può sorgere per via delle diverse funzioni svolte dall'intonazione, molte delle quali sono di tipo comunicativo e non solo strettamente linguistico (Kohler, 2006; Xu, 2006). Inoltre, non è chiaro se gli eventi intonativi siano percepiti categoricamente, caratteristica che faciliterebbe il fatto di riferirsi ad essi come ad eventi fonologici. Al contrario, gli eventi intonativi sembrano poter essere simili ai segmenti vocalici, ossia non sono caratterizzati da discontinuità di tipo articolatorio. Questo suggerisce che le soglie percettive degli eventi intonativi siano poco definite e che quindi gli eventi stessi possano non essere percepiti categoricamente.

In questo contributo si prenderanno in considerazione i risultati dell'analisi di dati di produzione e di percezione relativi all'intonazione della varietà pisana di italiano, come esempio di lingua in cui gli eventi tonali riguardano il livello post-lessicale (Gili Fivela, 2008). Si evidenzierà l'esistenza di eventi tonali che svolgono funzioni linguistiche specifiche, come gli elementi di una categoria dovrebbero fare, ma si sottolineerà anche l'eterogeneità delle caratteristiche di questi eventi tonali in alcuni contesti, sia dal punto di vista acustico che percettivo (e quindi interpretativo, dati i test utilizzati). Si discuterà della presenza di categorie intonative dai tratti peculiari, strettamente collegati alla specifica funzione svolta dall'intonazione nelle lingue non tonali: le categorie intonative saranno individuate come categorie linguisticamente rilevanti, pertinenti all'interno del sistema fonologico, benché non necessariamente discrete (per approfondimenti, Gili Fivela, 2008).

2.1 Gli eventi intonativi, le funzioni linguistiche e i significati

L'analisi di materiale letto e semi-spontaneo, raccolto registrando parlanti della varietà pisana di italiano[1], ha evidenziato l'esistenza di

[1] Si tratta di frasi molto controllate dal punto di vista della composizione segmentale e dell'interpretazione pragmatica e di

andamenti melodici specifici in associazione a funzioni e significati di tipo linguistico. Diversi accenti tonali e toni di confine[2] sono stati individuati in affermazioni, in domande polari e domande aperte, in contesti di focalizzazione e in corrispondenza di eventi di *phrasing* (cioè della segmentazione prosodica dell'enunciato). L'analisi dei materiali ha permesso di individuare elementi tonali caratterizzati da una forma fonetica specifica e facenti parte di andamenti nucleari (che riguardano le teste dei sintagmi prosodici) con funzione distintiva all'interno del sistema linguistico.

L'analisi dei materiali non ha portato alla luce una corrispondenza biunivoca tra *pattern* intonativi, ossia andamenti melodici, e funzioni o significati identificati grazie all'analisi delle mosse conversazionali (individuate sulla base del modello del discorso esplicitamente proposto per i dialoghi Map-Task - Carletta et al. 1995; 1996). Ad esempio, un accento discendente seguito da toni di confine bassi si è rivelato importante per la descrizione in termini fonologici sia di un contorno affermativo sia di un contorno interrogativo corrispondente ad una domanda aperta (i due contorni sono invece differenziati da caratteristiche fonetiche). Nonostante questo, l'identificazione di contesti specifici (di alcune funzioni e significati) associati con sequenze definite di eventi tonali ha permesso di prendere in considerazione questi eventi tonali come candidati per la definizione del sistema intonativo, come elementi costitutivi dei contorni intonativi associati a funzioni diverse nell'uso della lingua. Gli andamenti tonali sono stati quindi associati a specifiche funzioni; nel caso in cui singoli eventi dalle caratteristiche fonetiche simili fossero poi individuati in contesti differenti, non è stato necessario postulare l'esistenza di nuove unità costitutive: la differenziazione funzionale,

---

dialoghi registrati con il metodo del Map-Task (Anderson et al. 1991). In particolare, il Map-Task – che consiste nel ricostruire un tracciato su una mappa, grazie allo scambio verbale di informazioni - permette un controllo abbastanza elevato circa la struttura segmentale e soprasegmentale di alcune parole bersaglio, in corrispondenza delle quali saranno studiati gli eventi tonali; inoltre permette un buon controllo della funzione linguistica svolta dagli enunciati in cui le parole bersaglio sono utilizzate.

[2] Gli accenti tonali sono associati a sillabe, mentre i toni di confine sono associati a confini prosodici e, solo secondariamente, a sillabe.

ad esempio, è stata attribuita all'andamento nel suo complesso. In ogni caso, le unità costitutive postulate sono di tipo fonologico, ma a differenza di unità di livello segmentale, sono state associate a significati molto ampi (come introduzione, selezione, selezione ed opposizione rispetto ad altri elementi, secondo quanto già proposto per altre lingue – Gussenhoven, 1984; Pierrehumbert and Hirschber 1990, Kohler, 1991; 2006 – cfr. tab. 2), proprio sulla base dell'osservazione delle loro caratteristiche all'interno dei contesti individuati.

Nell'analisi dei dati di produzione, la situazione che più si è avvicinata all'esistenza di una relazione biunivoca tra *pattern* e funzione (situazione ideale che, come già detto, non è stata riscontrata) è quella in cui un *pattern* si è rivelato come specifico per una funzione, indipendentemente dal fatto che altri *pattern* possano essere stati osservati in relazione alla stessa funzione. Questo è avvenuto, ad esempio, nel caso delle domande polari, sia quelle formulate per la richiesta di nuove informazioni sia quelle volte alla richiesta di conferma. In questi casi, infatti, un andamento melodico tra quelli osservati nelle domande polari per richiesta di nuove informazioni (corrispondente a un accento nucleare discendente, seguito da un tono di confine alto-basso - secondo il quadro autosegmentale-metrico, H+L* H-L%) non è stato riscontrato nelle mosse instruct, explain e replay - tra gli enunciati affermativi - o in altri tipi di domanda; in modo analogo, un andamento specifico (corrispondente a un innalzamento-abbassamento nucleare, seguito da un innalzamento al confine – analizzato come H*+L L-H%) è stato individuato nelle domande effettuate per richiesta di conferma e nelle mosse *object*[3], ma non in altre mosse, ad esempio in affermazioni. In questi casi, quindi, si osserva la presenza di un *pattern* specifico che svolge una precisa funzione (ad es. H+L* H-L% per la richiesta di conferma), benché la stessa funzione possa essere svolta anche da altri andamenti melodici (ad es., [L+]H*+L L-L%). L'esistenza di un andamento 'dedicato' evidenzia quindi che domande polari, specifiche richieste di conferma

[3] Mosse conversazionali grazie alle quali il parlante manifesta incredulità rispetto alle informazioni fornitegli dall'interlocutore che dovrebbe svolgere il ruolo dominante nello scambio comunicativo, ossia l'*instruction giver* (poiché possiede le informazioni per alla ricostruzione del tracciato nel Map-Task).

e *object move* possono essere differenziate da altre mosse grazie alla sola intonazione.
In generale, però, si osservano corrispondenze non biunivoche, per quanto riguarda sia la relazione tra *pattern* e funzione sia la relazione tra funzione e *pattern* – si veda lo schema nella tab. 1. Diverse lingue, o varietà, possono presentare caratteristiche differenti rispetto al numero di *pattern* che si realizzano in corrispondenza delle varie funzioni. In ogni caso, osservazioni circa la mancanza di una relazione biunivoca tra andamento melodico e funzione possono aver favorito la posizione di alcuni autori che sostengono che l'intonazione non svolge un ruolo fonologico all'interno del sistema linguistico (Marotta, 2003, in press). Al contrario, l'estrema variabilità è qui intesa come correlata al fatto che, quando un enunciato è utilizzato per esprimere una certa funzione, veicola anche altri significati (Kohler, 2006; Xu, 2006), in misura variabile. Questo è un ruolo fondamentale della prosodia e dell'intonazione, reso possibile dalla presenza di eventi tonali che rappresentano gli elementi costitutivi dei contorni intonativi - in termini di forma e funzione – e che, di per sé, sono dotati di significati abbastanza ampi (cfr. § 2.1). In questa sede si propone che i significati ampi possano rendere possibili produzioni nelle quali si realizzano diverse sfumature di significato e nelle quali, grazie ad un processo composizionale in cui sono coinvolti anche i significati di altri eventi tonali, sono incluse tutte le informazioni volute dal parlante e che saranno poi disponibili in percezione, comprese quelle strettamente linguistiche. Questa caratteristica viene qui considerata alla base della flessibilità dell'intonazione e del suo essere funzionale a comunicare varie informazioni su un piano parallelo rispetto al livello lessicale, semantico, sintattico e pragmatico.

Function /move type	Nuclear Pattern
Wh-questions/ query-w	H+L*n L-L%
yes/no-questions query-yn	H+L*n H-(L%) [L+]H*n L- H+L* L-L% or L-H%
Check&align	(H+)L*n H-(L%) (check, align, and tag) [L+]H*+Ln L-L% or L-H% (check) [L+]H*n L-L% or H-H% (align and tag[4])
Object	[L+]H*+Ln[5] L-H% >[L+]H*+Ln L-H% or L-L%
questions (alternatives & elliptic)	[L+]H* (H+)L*n L-L% or H-H% (two alternat.) [L+]H*n L-L% (alternatives) [L+]H*n H-H% (elliptic)
statements instruct & explain & reply	H+L*n L-L% or H-H% [L+]H*n H-H% or L-L%
Focus	H+L* H+L* H- (wh- and yes/no-quest) [L+]H* H+L* (statement: narrow) [L+]H*+L (statement/check: narrow contrastive)
Phrasing	L- H- L-L%; H-H%; H-L%; L-H%

Tabella 1: Funzioni dei principali andamenti nucleari (da Gili Fivela, 2008)

Le diverse funzioni dell'intonazione (e della prosodia) possono influenzare il modo in cui gli eventi tonali (gli elementi costituenti della melodia, che potremmo definire "primitivi", ma anche i *pattern* nel loro complesso) sono selezionati e realizzati: influenzano il modo in cui sono selezionati perché in alcuni contesti i parlanti possono scegliere diversi eventi o *pattern* per esprimere una funzione; influenzano il modo in cui sono implementati perché in alcuni contesti i parlanti possono modificare l'implementazione fonetica di un evento tonale o di un *pattern*. Tra i fattori rilevanti nel modificare scelte e realizzazioni si riscontrano soprattutto gli aspetti paralinguistici, legati all'atteggiamento del parlante ed alla struttura del'informazione, sia a livello di frase che di discorso (tralasciando le emozioni - vedi Gili Fivela, 2008, in particolare § 1.2.3 e § 1.3.3). Ad esempio, i parlanti pisani chiamati a realizzare delle mosse *instruct* o delle domande

[4] Nelle *tag question* si riscontrano solo toni di confine H-H%.

[5] L'accento è spesso caratterizzato da allineamento ritardato e/o escursione tonale ampia, chiaramente visibile già dall'inizio dell'enunciato, etichettabile con %H.

polari, selezionano diversi eventi tonali sulla base dei fattori appena menzionati: selezionano toni di confine alti o bassi per esprimere la continuazione, accenti ascendenti o discendenti per esprimere la focalizzazione ristretta, o diversi *pattern* per esprimere le domande polari. Peraltro nel caso di accenti discendenti, i parlanti incrementano i valori di frequenza fondamentale del tono alto in un accento discendente (H+L*) quando vogliono essere più assertivi, così come cambiano l'intervallo di frequenza utilizzato per i sintagmi e l'altezza dei toni di confine per ragioni collegate alla struttura informativa.
D'altro canto, diverse funzioni possono essere svolte dallo stesso *pattern* o dalla stessa sequenza di eventi tonali. Ad esempio, l'andamento melodico riscontrato nelle domande aperte è caratterizzato dalla stessa sequenza di eventi tonali che si ritrova nelle mosse *instruct*, *explain* e *reply*, tra le affermazioni. La differenza tra le varie occorrenze può essere individuata nella realizzazione fonetica: i parlanti, infatti, possono realizzare diversamente elementi appartenenti alla stessa categoria, modificando il modo in cui esprimono una funzione, oppure possono realizzare diversamente gli elementi appartenenti alla stessa categoria per esprimere una funzione diversa. Ad esempio, possono aumentare il valore di frequenza fondamentale del *leading tone* alto in un accento come H+L*, per maggiore enfasi (si vedano anche i codici biologici di Gussenhoven, 2002; 2004), oppure possono modificare le caratteristiche fonetiche, ad esempio il *pitch range*, per differenziare affermazioni e domande; inoltre, come indicato da una serie di esperimenti percettivi, possono sfruttare le informazioni sullo *scaling* dei bersagli tonali per passare da un'interpretazione contrastiva ad una *broad focus* degli stessi accenti – Gili Fivela, 2006).
Osservando con attenzione i casi nei quali lo stesso *pattern* viene usato in associazione a diverse funzioni, emergono aspetti corrispondenti tra le funzioni in questione, oppure emerge il fatto che una possibile disambiguazione delle funzioni può essere garantita da altre considerazioni oltre a quelle fonetiche appena menzionate - ad esempio, di tipo lessicale, sintattico e pragmatico. Per quanto riguarda l'esistenza di una relazione tra le diverse funzioni espresse dallo stesso *pattern*, basti pensare all'andamento tipico della continuazione, riscontrato, dal punto di vista fonologico, sia in affermazioni che in domande ellittiche: di fatto, comunicare che c'è un elemento mancante che potrebbe seguire, può essere simile al formulare una domanda

ellittica, che presuppone che l'interlocutore, rispondendo, fornisca un'informazione mancante. Per fare un altro esempio, si pensi all'uso di un *pattern* ascendente-discendente-ascendente per le *object move* (cfr. nota 3 e tab. 1) e per le domande per richiesta di conferma prodotte dal parlante che svolge il ruolo dominante nella conversazione. In entrambi i casi, il contorno melodico è usato per opporsi al punto di vista dell'interlocutore, da parte di chi pensa di avere tutte le informazioni necessarie a sostenere, appunto, una diversa opinione; tuttavia la mancanza di intonazione assertiva (prevalentemente discendente), indica in ogni caso l'esigenza di una verifica. Per quanto riguarda i casi in cui una possibile disambiguazione circa la funzione ed il significato possa essere garantita da altri aspetti oltre a quelli fonetici, si pensi al fatto che la presenza vs assenza di un pronome interrogativo può rappresentare un'indicazione abbastanza chiara circa la differenza tra domande aperte e affermazioni, indipendentemente dalla presenza della stessa sequenza tonale (oltre a possibili differenze di *pitch range*); per fare un altro esempio, le 'formule' utilizzate nelle *tag question* (es., no?) sono un'indicazione chiara circa l'interpretazione da assegnare, benché la sequenza di toni che caratterizza queste domande possa essere analoga a quella individuabile in enunciati con funzione diversa, in cui non si riscontrerebbe rispetto allo stesso materiale lessicale (es., in mosse *instruct*, *explain* o *reply*).

Il quadro che emerge è di un insieme di entità tonali che solo in alcuni casi, in particolare in quelli in cui gli eventi codificano le sole informazioni disponibili per la corretta interpretazione della frase, appartengono ad andamenti che esprimono una sola funzione e corrispondono ad elementi realmente discreti (si veda l'interpretazione affermativa o interrogativa, legata alla presenza di un tono di confine basso o alto, L- o H-, come è stato anche confermato dagli esperimenti di percezione categorica; cfr. § 3, Gili Fivela, 2008, in prep). Negli altri casi, le indicazioni fornite dall'intonazione possono essere considerate come parallele rispetto a quelle provenienti da altri settori della grammatica. Pur essendo codificate in parallelo, però, rappresentano informazioni linguistiche oltre che paralinguistiche (ad es., indicazioni circa la focalizzazione, che può cambiare il valore di verità di un enunciato – Rooth, 1985, o circa il *phrasing*, che permette di disambiguare informazioni legate alla struttura sintattica – come nel caso dell'*attachment* di sintagmi preposizionali). La possibilità di una

codifica parallela fornisce probabilmente la libertà di scelta che permette ai parlanti di adattare al meglio la propria produzione al contesto specifico. Ad esempio, i pisani possono scegliere di usare sia H* che H*+L per esprimere un forma di contrasto (cfr. tab 1), a seconda che il contrasto sia sul piano sintagmatico o paradigmatico. Pensando ad un contesto in cui sia necessaria una correzione, i parlanti possono produrre un accento discendente (H+L*) piuttosto che un accento ascendente-discendente (H*+L), realizzando una correzione più neutra o meno enfatica. La presenza di ulteriori informazioni relative all'interpretazione dell'enunciato, oltre a quelle intonative, ad esempio può garantire la comprensione del messaggio di base: se un'affermazione è preceduta da un chiaro segnale di dissenso, come 'no', le informazioni intonative possono essere variate (e possono corrispondere, ad esempio sia ad un accento H+L* che ad uno H*+L) senza che possano esserci problemi in relazione al fine comunicativo. Questo non significa che l'informazione fornita dall'intonazione non sia di tipo linguistico o che sia sempre possibile individuare informazioni che rendano quelle prosodiche, in un certo senso, ridondanti. Piuttosto significa che, a parte i casi in cui l'intonazione rappresenta l'unico correlato per la disambiguazione del significato e della funzione (si veda l'andamento discendente-ascendente-discendente che è stato trovato nelle domande e non nelle affermazioni), l'intonazione può rappresentare un canale di informazione parallelo che, benché in una certa misura possa essere ridondante rispetto al significato di base, può invece essere primario per l'interpretazione completa, e talvolta corretta, dell'enunciato (influenzata anche dalla forza illocutiva e dalla struttura informativa dell'enunciato, oltre che da aspetti sociolinguistici, come suggerito da Marotta e colleghi - Marotta et al. 2004; Marotta, 2005 e lavori successivi). Per una discussione dell'influenza di questi fattori Gili Fivela (2008, soprattutto § 5.1.1); in questa sede è importante ricordare che si tratta, come già menzionato, di influenze sia circa la scelta degli eventi da realizzare (tra quelli presenti nell'inventario fonologico), sia delle loro caratteristiche fonetiche.

2.1 Significato e sfumature di significato delle componenti tonali

A parte le differenze riscontrate nella composizione degli andamenti melodici associati alle diverse funzioni, è emerso chiaramente che è

molto limitato il numero di elementi di base, ossia il numero di eventi tonali che compongono gli andamenti e che specificano composizionalmente le loro interpretazioni. Questi elementi sono stati considerati al fine di individuare il loro significato specifico, quello che li caratterizza in tutti i contesti nei quali compaiono. L'analisi ha fornito i risultati riportati nella tab. 2. I significati sono risultati molto ampi, analogamente a quanto già riportato in letteratura (Gussenhoven, 1984; Pierrehumbert and Hirschber 1990; Kohler, 1991, 2006). In ogni caso, poiché abbiamo visto che gli eventi tonali possono essere realizzati con estrema variabilità, non dovrebbe sorprendere l'impossibilità di riscontrare un elevato numero di usi contrastivi degli eventi intonativi di base (o volendo 'primitivi') individuati.

Secondo la posizione qui sostenuta, a seconda del significato e della sfumatura di significato che vuole veicolare, il parlante può effettuare scelte in relazione agli eventi tonali da realizzare o in relazione alla loro implementazione. Le differenze riscontrate nella composizione del *pattern*, ma soprattutto nella sua realizzazione, possono portare ad uno slittamento di significato, per il quale un evento tonale può avere un'interpretazione simile a quella di un altro evento tonale (cfr. §4, fig. 1).

Function / meaning	Tonal event
Introduction in the discourse/shared assumptions conveying finality and/or self-confidence	H+L*
Selection from the discourse/shared assumptions: • for orienting attention • for (re)introducing • for adding related information - conveying continuation	H*
Opposition to the discourse/shared assumptions Conveying finality and/or self-confidence	H*+L
Ending of a unit • global end • for adding more material – partial end	L- and L%
Continuation of a unit • by the same speaker • by the interlocutor	H- and H%

Tabella 2: Significati e funzioni degli eventi tonali individuati (da Gili Fivela, 2008)

La conseguenza immediata è che le categorie intonative possono non essere discrete per quanto riguarda il riferimento al significato e la loro percezione, o comunque possono non essere discrete nei termini in cui lo sono altre unità linguistiche o nei termini in cui si sarebbe tentati di considerarle dal momento che sono intese come fonologiche (cioè discrete nel senso di passibili di percezione categorica). In secondo luogo, non sorprende che nei dati di produzione le proprietà fisiche degli eventi tonali varino all'interno di una sorta di 'area di esistenza'. Diversamente da quanto accade nel dominio segmentale, la variazione all'interno di questa area di esistenza corrisponde ad informazioni specifiche: la realizzazione in una zona dell'area di esistenza, ad esempio della categoria A, che sia vicina al confine con un'altra area, ad esempio B, corrisponde ad uno slittamento di significato, ad una sfumatura di significato e ad una funzione di A che siano vicine al significato e alla funzione normalmente associate con l'altra area, quella di B (cfr. § 4, fig. 1).
Questa variazione è collegata alla necessità di codificare diversi tipi di informazione e, soprattutto quando l'intonazione non rappresenti il solo correlato di una differenza interpretativa, può interferire con la capacità di percepire categoricamente le informazioni derivanti dall'intonazione. Come si vedrà nel prossimo paragrafo, infatti, i risultati dei test percettivi mostrano che le informazioni intonative sono utilizzate ma, in molti casi, non tramite percezione categorica.

3 La percezione degli eventi tonali

Al fine di verificare la realtà percettiva degli accenti tonali e dei toni di confine individuati nei materiali di produzione e di appurare se si tratti di eventi discreti, sono stati organizzati test di percezione categorica in cui gli eventi indagati sono stati considerati a coppie nel contesto tonale più simile possibile (Gili Fivela, 2008)[6]. I risultati

[6] Il paradigma è quello proposto in Lieberman et al. (1957). La percezione si considera categorica se i soggetti riescono ad identificare e discriminare stimoli che siano stati manipolati acusticamente in modo da rappresentare un *continuum* di variazione tra due categorie. I soggetti devono identificare elementi appartenenti a due categorie diverse, in modo che le loro risposte dimostrino un passaggio brusco, e non graduale, nell'identificazione di una o

mostrano che mentre si riscontra percezione categorica nel caso di accento di confine intermedio alto vs basso - nel contesto della distinzione tra domande polari e affermazioni - non si individua percezione categorica nel caso dell'accento ascendente vs ascendente-discendente e dell'accento discendente vs ascendente-discendente - che esprimono focus ristretto, focus contrastivo e focus ampio. Di fatto i soggetti che hanno partecipato agli esperimenti sono stati in grado di interpretare categoricamente gli stimoli (in particolare identificando chiaramente quelli agli estremi del *continuum*), ma non attraverso percezione categorica (ossia senza che un picco di discriminazione in corrispondenza degli stimoli di transizione tra le due categorie, nella parte centrale del *continuum*).
Nello studio relativo agli accenti tonali sono stati variati separatamente lo *scaling* e l'allineamento dei bersagli tonali, creando un *continuum* di stimoli che, a parte gli estremi, erano sempre ambigui per le caratteristiche di almeno un correlato. I risultati hanno mostrato che gli estremi di manipolazione sono stati identificati come appartenenti a diversi insiemi, in alcuni casi a seconda della manipolazione sia dell'allineamento che dello *scaling* (si veda il confronto tra accento ascendente-discendente vs ascendente), in altri casi a seconda della variazione del solo *scaling* (si veda il confronto tra accento discendente vs ascendente-discendente, almeno per gli stimoli creati a partire dall'enunciato base contrastivo). Risultati di un recente studio (Gili Fivela, in prep.), effettuato sulle stesse coppie di accenti manipolando coerentemente allineamento e *scaling*, hanno confermato l'interpretazione categorica degli stimoli all'estremo del *continuum* ma l'assenza di percezione categorica. I risultati mostrano che l'interpretazione categorica può avvenire prestando attenzione a diversi correlati – qui, allineamento e *scaling* – o ad uno solo – ad esempio l'allineamento (con differenti strategie a seconda del soggetto). Inoltre, soprattutto nel confronto tra accento ascendente-

---

dell'altra categoria. Inoltre, nel momento in cui si richieda ai soggetti di discriminare coppie di stimoli adiacenti nel *continuum* di variazione fonetica, ci si aspetta che siano in grado di discriminare solo gli stimoli di transizione tra le due categorie, quelli corrispondenti a circa il 50% di identificazione. Di fatto ci si aspetta che i membri di una categoria presentino caratteristiche omogenee e siano discriminabili solo gli elementi di transizione tra categorie.

discendente e ascendente, lo *scaling* influisce chiaramente sulla percezione delle categorie, determinando scarsi livelli di percezione di una categoria quando i suoi valori siano coerenti con la categoria opposta (quindi uno slittamento verso la categoria opposta nella fase di identificazione). Di fatto, comunque, non si può parlare in alcun caso di percezione categorica quando si considerino gli accenti tonali, ma solo di interpretazione categorica (Ladd and Morton, 1997). Al contrario, la percezione è risultata categorica nel caso degli eventi di confine legati alla differenziazione di domande e affermazioni: gli stimoli sono stati infatti identificati come appartenenti a due categorie diverse, con un passaggio brusco nei giudizi di appartenenza all'una e all'altra; inoltre gli stimoli corrispondenti alla transizione tra le due categorie sono stati chiaramente discriminati.
Nel tentativo di trovare evidenza circa la discretezza delle categorie fonologiche, in modo da poter sostenere agevolmente che svolgano un ruolo linguistico, è stato anche svolto un compito di imitazione degli stessi *continua* di stimoli[7]. Per quanto i soggetti che hanno partecipato agli esperimenti di imitazione siano pochi, i risultati indicano che essi possono imitare gli stimoli come se appartenessero a insiemi discreti, supportando l'ipotesi dell'esistenza di categorie distinte. Tuttavia i risultati mostrano anche una differenza a seconda dell'evento tonale considerato e quindi della categoria: i soggetti imitano il *continuum* di stimoli realizzando due *pattern* discreti nel caso del *continuum* tra un accento ascendente ed uno ascendente-discendente, mentre la discretezza è meno evidente nel caso dell'imitazione del *continuum* tra un accento discendente ed uno ascendente-discendente. D'altro canto i soggetti producono *pattern* chiaramente discreti nel caso della presenza di un tono di confine alto ed uno basso, in relazione alla differenza tra affermative e domande polari. Quindi l'eterogeneità degli eventi considerati si riscontra sia nel test di percezione categorica che in quello di imitazione, benché i parlanti usino in

[7] Nel caso del test di imitazione (Pierrehumbert and Steele, 1989), si chiede ai soggetti di imitare gli stimoli che rappresentano il *continuum* di variazione per verificare se, nella produzione, realizzino elementi dalle caratteristiche chiaramente ascrivibili a due categorie distinte (piuttosto che riprodurre il *continuum* di variazione dall'una all'altra – cfr. Gili Fivela, 2008 per una discussione più approfondita).

produzione tutti gli elementi tonali testati, associando ad essi diverse funzioni e significati.
Se dovessimo considerare come fonologici solo gli eventi per cui si sono ottenuti risultati coerenti in tutti i test, dovremmo sostenere che solo il tono di confine alto e basso, che differenziano le domande polari dalle affermazioni, appartengono al sistema fonologico. Questo comunque corrisponderebbe a considerare come una distinzione dicotomica quella che sembra piuttosto essere un distinzione in termini di grado tra risultati più o meno simili a quelli attesi per la presenza di categorie linguistiche (ad un estremo, infatti, abbiamo sia percezione categoriale che capacità di imitazione del *pattern* e all'altro estremo abbiamo l'assenza di percezione categorica e diversi gradi di accuratezza nell'imitazione).
L'ipotesi qui sostenuta è che i significati e le funzioni associate agli eventi tonali indagati possano essere rilevanti nello spiegare i diversi risultati dei test percettivi. La mancanza di percezione categorica può essere spiegata con riferimento al fatto che, in molti dei casi qui considerati, i *pattern* tonali sono effettivamente non discreti, prima di tutto a causa dei significati e delle funzioni che ricoprono: funzioni e significati, poi, rispetto ai quali l'intonazione rappresenta informazione aggiuntiva rispetto a quella di base contenuta nell'enunciato (ad esempio, il contenuto lessicale e semantico del sintagma potrebbe già essere sufficiente a correggere un'affermazione precedente). Tuttavia, l'interpretazione categorica degli stimoli agli estremi del *continuum* di manipolazione, insieme alla capacità di produrre eventi tonali discreti imitando un *continuum* di variazione mostra l'esistenza di categorie.
Inoltre i soggetti potrebbero riuscire più facilmente a produrre elementi discreti piuttosto che a percepire in modo categorico gli eventi (come se fossero discreti), in modo analogo a quanto osservato da Labov (1994), e si può ipotizzare che questo dipenda da un'asimmetria tra produzione (in questa sede, nel parlato letto, semi-spontaneo e in quello ottenuto nei test di imitazione) e percezione (in questa sede, la percezione categorica). Peraltro sono numerosi i lavori che evidenziano le differenze tra percezione e produzione, benché talvolta con esiti discordanti. Ad esempio Labov (1994) riporta che i parlanti sono in grado di produrre ciò che non riescono a percepire, mentre Warren et al. (in press) riscontrano che i soggetti percepiscono differenze che non sono in grado di produrre. Sulla base della

discussione precedente, però, si potrebbe anche ipotizzare che i soggetti percepiscano il *continuum* di variazione (ossia differenze interne alle categorie) perché solitamente vi prestano attenzione, visto che grazie alle variazioni ricevono informazioni rilevanti (infatti, i risultati di percezione mostrano un cambiamento abbastanza graduale nell'identificazione dei *pattern*). D'altro canto, quando i soggetti siano forzati a realizzare un evento tonale o un *pattern*, codificano le informazioni importanti per veicolare il messaggio: quando devono esclusivamente imitare un enunciato possono non riuscire ad identificare informazioni rilevanti a parte il significato principale o la funzione principale espressi dal *pattern* scelto. Non a caso, il compito di imitazione è stato proposto da Pierrehumbert and Steele (1989) proprio ipotizzando che i parlanti producano degli andamenti intonativi corrispondenti alle categorie rappresentate dagli estremi del *continuum*, senza modificazioni graduali. In ogni caso, la complessità dei risultati ottenuti nei test di percezione categorica, unita al fatto che gli eventi tonali sono chiaramente differenziati, almeno per via di come i parlanti li usano nei materiali di produzione, suggerisce che il compito di percezione categorica non sia ideale per verificare il ruolo linguistico delle informazioni intonative; infatti l'interpretazione categorica degli stimoli agli estremi del *continuum* può essere sufficiente a garantire l'esistenza di più eventi intonativi e a mostrare che i soggetti usano le informazioni per interpretare il messaggio, indipendentemente dal fatto che percepiscano categoricamente i cambiamenti intonativi.

Al fine di ottenere ulteriori dati circa l'esistenza di vere e proprie categorie intonative, che forniscano informazioni linguistiche, è stato effettuato un test ulteriore, prendendo in considerazione la coppia di accenti intonativi rispetto alla quale si erano ottenuti i risultati meno chiari. In questo caso, ai soggetti era richiesto di spostare due carte su un tavolo, sulla base di una sequenza di tre istruzioni, di cui la centrale poteva corrispondere ad una correzione rispetto all'istruzione precedente (questo test è stato chiamato card game task - Gili Fivela, 2008). I risultati hanno mostrato che i soggetti correggono la loro mossa precedente nei casi in cui lo stimolo si trovava all'estremo di variazione corrispondente all'andamento contrastivo. In ogni caso, i soggetti non si comportano in modo coerente con la presenza di percezione categorica (di fatto non è stato individuato un brusco cambiamento delle azioni di risposta in corrispondenza della parte

centrale del *continuum* di stimoli), ma è chiaro che utilizzano le informazioni legate alla variazione intonativa. Quindi i soggetti percepiscono il *continuum* di variazione perché normalmente ricevono informazioni da tale *continuum*. Tuttavia essi usano il *continuum* per prendere decisioni categoriche circa il gioco di carte, correggendo o meno la loro mossa precedente.
Quindi, come già sottolineato da Kohler (2006), la percezione di categorie può essere considerata più importante della percezione categorica. In questo senso, la capacità dei soggetti di associare gli stimoli agli estremi del *continuum* a due diversi insiemi di funzioni, così come il loro uso delle informazioni percepite per prendere diverse decisioni rispetto a delle azioni, si considera sufficiente a mostrare che sono state percepite due categorie intonative diverse.

4 Sfumature di significato, codifica di più informazioni, ed 'area di esistenza' degli eventi tonali

Molti dei significati e delle funzioni espresse dall'intonazione sono di tipo linguistico, ad esempio si riferiscono alla modalità della frase, così come al focus e al *phrasing*. L'intonazione sembra offrire informazioni sul messaggio che possono essere parallele rispetto a quelle offerte da altri elementi del sistema linguistico. Tuttavia, come i test percettivi hanno mostrato, il ruolo delle modificazioni intonative non è omogeneo perché da un lato può rappresentare l'unico correlato per una interpretazione completa e corretta dell'enunciato (cfr. affermazioni vs domande polari), dall'altro può "solo" aggiungere informazioni rispetto a quelle codificate grazie agli altri livelli della grammatica. Inoltre, il parlante sembra poter esprimere le varie funzioni e i vari significati con diversi gradi di intensità, o arricchendole di diverse sfumature, sia modificando la composizione tonale del *pattern* intonativo, sia modificando la sua implementazione (le istruzioni possono essere più o meno assertive e perentorie, una richiesta di conferma può derivare da diversi gradi di sicurezza circa il presupposto di conoscenza o, per fare un ulteriore esempio, la focalizzazione può essere realizzata in modo più o meno enfatico). Quindi i significati e le funzioni sembrano essere caratterizzati da un 'nucleo' - un significato o una funzione primaria, prototipica - e da possibili differenziazioni che ne rappresentano delle sfumature. Coerentemente, spesso non si osserva una relazione biunivoca tra

significato o funzione ed evento tonale o *pattern* melodico (cfr. tab. 1). In ogni caso, in base ai risultati dell'analisi dei dati di produzione e di percezione, così come in base ai lavori descritti in letteratura, gli eventi tonali che compongono i diversi *pattern* melodici risultano dotati di un loro significato ampio e di proprietà che partecipano alla codifica delle diverse funzioni. Nell'analisi del pisano (Gili Fivela, 2008), il significato degli eventi tonali è stato delineato facendo riferimento alla struttura dell'informazione – introduzione, selezione di un elemento e sua opposizione ad altri elementi nel discorso – ed è stato caratterizzato nei termini di un insieme di componenti, identificate come tratti/proprietà condivise dai *pattern* in cui un evento tonale è stato individuato (cfr. tab.2).

Inoltre, considerando le componenti tonali degli eventi intonativi 'primitivi', sono state individuate alcune caratteristiche coerenti con le grammaticalizzazioni proposte in relazione all'esistenza di codici biologici (Gussenhoven, 2002; 2004; cfr. Gili Fivela, 2008, §3.4). Si veda la figura 1 per una rappresentazione schematica dei significati associati agli eventi tonali nell'italiano di Pisa e delle direzioni di variazione in base alle loro componenti tonali. In sintesi, è stato proposto che i significati e le funzioni degli eventi tonali siano corrispondenti a componenti che sono anche coerenti con le caratteristiche fonetiche della composizione tonale degli eventi stessi, in termini di specificazioni tonali alte e basse. Seguendo la proposta di Gussenhoven (2002; 2004), ad esempio, sono state prese in esame le componenti di 'finalità' o 'sicurezza' in relazione alla presenza di specificazioni tonali basse, corrispondenti a valori bassi o discendenti; analogamente, le componenti di 'continuità' o 'incertezza' sono state correlate alla presenza di specificazioni tonali alte, corrispondenti a valori alti o ascendenti di frequenza fondamentale (coerentemente con i codici di frequenza e di produzione elaborati da Gussenhoven ibidem).

Le differenze osservate nei *pattern* tonali, sia nei termini di scelta degli eventi tonali che di loro realizzazione, possono quindi essere viste come correlate a cambiamenti legati alle componenti di significato e, di conseguenza, al significato che gli eventi tonali veicolano. Modificando i correlati (o uno dei correlati) di un evento tonale si può ottenere uno slittamento del significato associato nella direzione di un significato corrispondente ad una forma fonetica simile a quella dell'evento tonale realizzato (e modificato). Infatti la

variazione in allineamento e *scaling* ha spesso svolto un ruolo importante, nei dati analizzati, nel codificare diverse sfumature di significato, facendo slittare i significati dei *pattern* verso significati veicolati principalmente da altri eventi tonali, e facendo acquisire loro le relative componenti di significato (si pensi ai risultati dei test percettivi per gli accenti tonali). Quindi la realizzazione degli eventi tonali, così come la percezione dei significati da loro veicolati, potrebbe avere luogo in una sorta di 'area di esistenza' nella quale le differenze siano informative rispetto al messaggio. La situazione relativa all'italiano di Pisa può essere rappresentata schematicamente come nella figura 1, dove sono mostrate anche le possibili intersezioni tra le componenti[8].

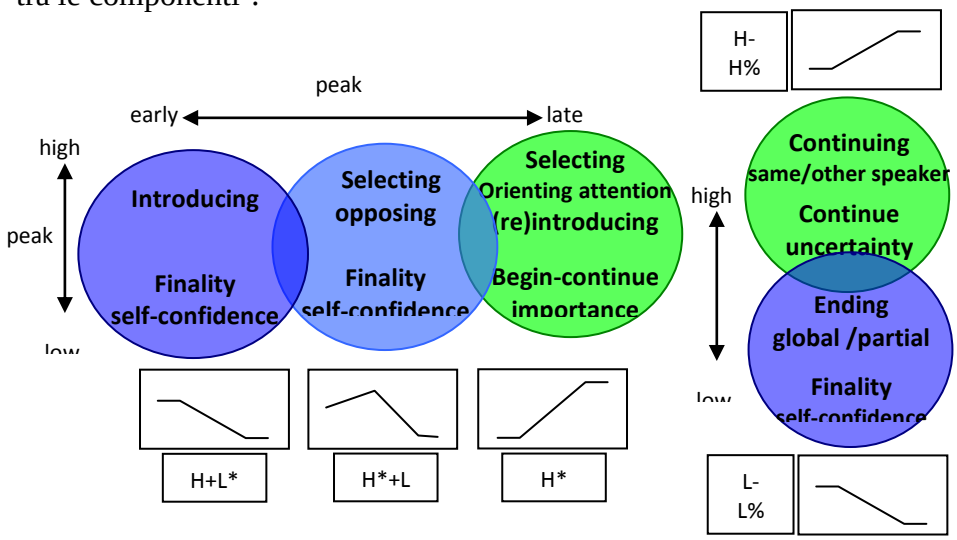


Figura 1: Rappresentazione schematica dei significati e dell'intersezione delle componenti tonali nell'italiano di Pisa; nello schema è anche indicata la direzione di variazione in base alla variazione di frequenza (da Gili Fivela, 2008).

Ad esempio, realizzando un accento tipo H*+L piuttosto che un H+L* il parlante sceglie di veicolare chiaramente il significato di 'opposing'

[8] In questa fase, l'interazione con durata ed intensità non è esplicitamente considerata. Tuttavia, sulla base dei dati disponibili per il pisano, almeno per quanto riguarda gli accenti con componente bassa/discendente, si potrebbe ipotizzare un incremento dei valori di durata in relazione allo slittamento da un picco anticipato ad uno ritardato (ossia da H+L* a H*+L).

piuttosto che quello di 'introducing an element'; peraltro, producendo un accento come H+L* con un *leading* tone alto realizzato a frequenze particolarmente elevate, il parlante mantiene la funzione di introduzione dell'elemento, ma fa slittare l'interpretazione verso quella dell'accento H*+L (caratterizzato da un picco ritardato e in molti casi più elevato in termini di frequenza, cfr. Gussenhoven, 2004): H+L* assume parte dell'interpretazione di H*+L nei termini di 'selecting-opposing', pur mantenendo le (condivise) componenti di 'finality' e 'high confidence'.
Per fare un altro esempio, spostando in avanti ed innalzando il valore del picco di frequenza di un accento come H*+L si indebolisce la componente di 'finality' e 'self-confidence' per incrementare quella di 'continuity', 'importance' e 'uncertainty' (coerentemente, ad esempio, un ritardo nell'allineamento del picco si osserva nelle mosse *object* - cfr. tab. 1). In modo analogo, quando l'accento H*+L è seguito da una specificazione tonale low-high, piuttosto che dalla usuale marca tonale bassa riscontrata nei contorni contrastivi, il *pattern* vede inibita la sua componente di 'finality' e 'self-confidence' per acquisire, in questo caso in veri e propri termini composizionali, le componenti di 'continuing' e 'uncertainty', veicolate dalla presenza della specificazione alta di confine. Non è quindi un caso che questo sia il *pattern* utilizzato per chiedere informazioni da parte degli *information giver*, che giocano un ruolo dominante nella conversazione: se usassero un normale *pattern* contrastivo, come invece fanno gli *information follower* (cfr. Gili Fivela, 2008), sembrerebbero certi di ciò che affermano e le loro produzioni non verrebbero interpretate come delle richieste di conferma, ma come delle affermazioni.

5 Riflessioni conclusive

Pensare ad eventi tonali che possono essere realizzati in modo variabile all'interno di un'area di esistenza permette di rendere conto del fatto che nel caso delle categorie intonative la variazione è attesa ed informativa. La variazione in allineamento e/o *scaling* può veicolare diverse sfumature di significato per eventi tonali la cui realizzazione è legata a funzioni differenti e influenzata da vari fattori, tra i quali, ad esempio, la struttura informativa dell'enunciato e del discorso – per una discussione dettagliata, si veda Gili Fivela (2008).
Questa proposta permette di rendere conto della variazione osservata nell'intonazione per la codifica delle informazioni e della mancanza di percezione categorica osservata dal punto di vista percettivo per la maggior

parte delle specificazioni tonali (si ricorda che fa eccezione l'opposizione legata al cambiamento di modalità della frase). Infine, la proposta riconosce l'esistenza di categorie, le cui possibili realizzazioni appartengono ad un insieme. Tra le realizzazioni, alcune saranno prototipiche, o vicine al significato centrale; altre saranno meno prototipiche e quindi periferiche rispetto al significato centrale e potenzialmente più vicine ad un altro significato o ad una sua componente. Queste considerazioni richiamano la 'exemplar theory' (Pierrehumbert, 2001; 2002) o comunque una visione della categoria intonativa come di un insieme costituito da elementi più o meno rappresentativi, nel quale possa essere individuato anche un prototipo. E' quindi possibile che, per verificare l'esistenza di categorie intonative, un test che prenda in considerazione insiemi non omogenei e prototipi possa dare risultati più chiari rispetto ad un esperimento di percezione categorica (per i primi risultati di verifica dell'esistenza di prototipi per le categorie intonative, si rimanda a Gili Fivela, in prep.).

In conclusione, i materiali studiati sino ad ora hanno suggerito che nel caso dell'intonazione le differenze di distribuzione delle realizzazioni degli eventi tonali siano legate ad informazioni importanti. Queste informazioni sembrano svolgere un ruolo rilevante nel far sì che le categorie non siano necessariamente omogenee e quindi possano anche essere non discrete.

Sia i dati di produzione che i dati di percezione mostrano infatti l'esistenza di categorie intonative, espressione di informazioni linguistiche utilizzate dai parlanti. Tuttavia le categorie intonative appaiano più flessibili di quanto solitamente si ritiene siano le categorie linguistiche di tipo fonologico: l'intonazione fornisce difficilmente informazioni relative a categorie discrete, proprio perché si compone di categorie altamente informative.

Riferimenti bibliografici

Clements, G.N. and R. Ridouane (2006) Quantal phonetics and distinctive features: a Review, in A. Botinis (ed.), *Proc. of the Workshop on Experim. Linguistics* Athens: Univ. of Athens, pp.17-24.

Anderson A., Bader M., Bard E. et al. (1991) 'The HCRC map task corpus', *Language and Speech*, 34, 4, pp. 351-366.

Bruce, G. (1977) *Swedish word accents in sentence perspective*. Gleerup: CWK.

Carletta, J. et al. (1996) HCRC Dialogue structure coding manual, in *Human Communication Research Center, Technical Report 82*, University of Edinburgh, June 96, pp. 1-27.

Fry, D., A. Abramson, P. Eimas and A. Liberman (1962) 'The identification and discrimination of synthetic vowels'. *L.&Sp.*, 171-89.

Gili Fivela, B. (2006) 'The coding of target alignment and scaling in pitch accent transcription', *Italian Journal of Linguistics*, 18.1, pp.189-221.

– (2008) *Intonation in Production and Perception: The Case of Pisa Italian*. Alessandria: Edizioni dell'Orso.

– (in prep.) 'Meanings, shades of meanings and prototypes of intonational categories', *Special issue on Prosody and meaning*, P. Prieto and G. Elordieta (eds.).

Gussenhoven, C. (1984) *On the grammar and semantics of sentence accents*. Dordrecht: Foris.

Kohler, K. (1991) Terminal intonation patterns in single-accent utterances of German: Phonetics, phonology and semantics, in *Arbeitsberichte 25* (AIPUK 25), University of Kiel, pp. 117-185.

– (2006) Paradigms in Experimental Prosodic Analysis: From measurements to function, in Sudhoff et al. (eds), *Methods in Empirical Prosody Research*, Berlin: W. de Gruyter, pp. 123-152.

Labov, W. (1994) *Principles of linguistic change: Internal factors*. Oxford: Oxford Blackwell, v. 1.

Marotta, G. (2003) L'illusione prosodica, in *Atti del convegno di studi in memoria di Tristano Bolelli*, Pisa, Italy. Novembre 28-29, 2003, G. Marotta (ed.), *Studi e saggi linguistici*, XL-XLI, pp. 237-258.

Pierrehumbert J. (1980) *The phonology and phonetics of English intonation*. PhD thesis, MIT – published in 1988 by IULC.

– (2001) Exemplar Dynamics: Word Frequency, Lenition and Contrast, in J. Bybee and P. Hopper (eds) *Frequency Effects and Emergent Grammar*, Amsterd.: John Benjamins, pp. 137-157.

– (2002) Word-specific phonetics, in C. Gussenhoven and N. Warner (eds) *Laboratory Phonology VII*, Berlin: Mouton De Gruyter, pp. 101-140.

Pierrehumbert J. and J. Hirschberg (1990) The meaning of intonational contours in the interpretation of discourse, in P.R. Cohen et al. (eds.), *Intentions in communication*, MIT Press, pp. 271-310.

Rosch, E. H. (1975) 'Cognitive Representations of Semantic Categories', *Journal of Exp. Psyc.: General*, 104, 3, pp.192-233.

Stevens, K.N. (1989) 'On the quantal nature of speech'. *J. Phon.* 17, pp. 3–45.

Stevens K.N. and S.J. Keyser (2010) 'Quantal theory, enhancement and overlap', *Journal of Phonetics*, 38, 1, pp.10-19.

Xu Y. (2006) Speech prosody as articulated communicative functions, *Speech Prosody* 2006, Dresden, Germany, pp. 218-221.

Warren P., Hay J. and B. Thomas (in press) The loci of sound change effects in perception and production, in J. and Cole Hualde J. (eds) *Papers in Laboratory Phonology* IX: *Change in Phonology*, Oxford: OUP.

Lingua parlata – lingua cantata

Cantabilità delle lingue: rivisitazione di un topos – analisi storico-culturale, dimensione linguistico-tipologica e analisi linguistico-musicologica

Elmar Schafroth
Romanische Sprachwissenschaft, Universität Düsseldorf
schafroth@phil-fak.uni-duesseldorf.de

Abstract

This article deals with the "singability" of languages. This subject will be analyzed on three levels of examinations. First of all the historically well-known topos of judging (spoken) languages with respect to their esthetical image will be amply discussed by giving a whole series of examples from the 16th, 17th and the 18th century and by quoting famous statements from composers like Mozart and writers like Rousseau about the quality of the French and Italian language when they are sung. Parting form Rousseau's linguistic characteristics of singability and nonsingability, the second part focuses on criteria that will allow a linguistic-typological description of languages with regard to their property to contribute to a high(er) degree of singability (phonetic and phonological features like vowel inventory and vowel distribution, accent, rhythmic patterns, syllable structure, sandhi phenomena). The last chapter points out how the Düsseldorf project data bank works and what kind of conclusions can be drawn form the meticulous analysis of airs being part of one and the same opera, but sung in different languages (for the present limited to French, Italian and German). Each opera – The Marriage of Figaro, The Magic Flute, Iphigenia in Tauris *and* Carmen – *will be examined in its original version and its translated or adapted version (hence on the basis of four recordings and four librettos). The results of this linguistic and musicological analysis will reveal for each language different degrees of suitability to be sung. Further studies might also take into account other languages like English, Spanish or Russian.*

1 Introduzione

In questo contributo mi dedicherò ad un campo interdisciplinare che tocca la linguistica, la musicologia e la storia culturale: le differenze prosodiche tra il parlare e il cantare, le conseguenze per la percezione uditiva e, all'occorrenza, la valutazione estetica (se le parole cantate divergono prosodicamente troppo dalla realizzazione parlata può

risultarne un suono d'insieme sentito come distorto, anomalo, per es. quando le regole dell'accento fonico di una lingua non vengono rispettate in una canzone).
Si tratterà di presentare un metodo linguistico per analizzare e descrivere la "cantabilità" di lingue diverse (qui l'italiano, il tedesco e il francese). Va da sé che ogni lingua può essere cantata – le tradizioni della musica popolare in tutto il mondo ne sono la prova –, ma ci sono delle differenze considerevoli tra le lingue se il testo (per esempio di un libretto) deve adeguarsi ad una musica (a diversi schemi ritmici, tempi, intervalli ecc.) già prestabilita. La "flessibilità" di una lingua di regolarsi alle esigenze delle condizioni ritmiche della musica si basa sulla sua "dotazione materiale", tra cui l'inventario di fonemi e la distribuzione fonemica, le strutture fonotattiche e prosodiche (sillabe, quantità vocalica e consonantica, schemi accentuativi e ritmici), fenomeni di variabilità (la possibilità dell'elisione di una vocale o di una sillaba, come il troncamento in italiano, o il contrario, cioè l'epentesi (anaptissi o paragoge). Il catalogo di questi tratti conduce ad una tipologia di parametri adatta a descrivere la flessibilità musicale di una lingua. Queste riflessioni fanno parte di un progetto di ricerca ("Lingua e musica: Analisi della prosodia di lingue cantate tramite arie d'opere e i loro libretti") che comprende tre parti: 1. "Analisi storico-culturale" (considerazioni sull'estetica delle lingue parlate e/o cantate, riferite già parzialmente all'antichità e approfondite dal Seicento in poi); 2. "Analisi linguistico-tipologica"; 3. "Analisi musicologico-linguistica" (analisi contrastiva dei libretti e delle arie d'opera, rilevata in una banca dati).

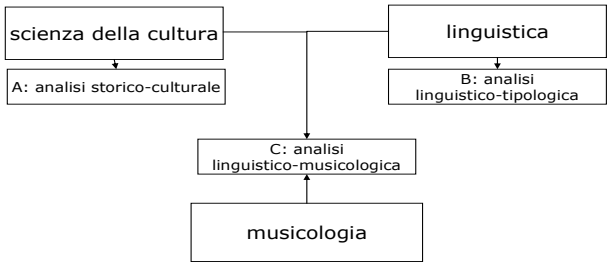


Figura 1: Approccio integrale per l'analisi del nesso tra lingua e musica

Nella mia comunicazione presenterò brevemente il progetto intero ponendo l'accento però sulla seconda parte (le differenze prosodiche tra lingua parlata e cantata).

2 Analisi storico-culturale

Un'analisi della "cantabilità" delle lingue basata su criteri linguistico-tipologici e linguistico-musicologici deve essere separata da giudizi stereotipici sul suono e sull'estetica delle lingue parlate o cantate. È probabile che, a seconda vista, in queste valutazioni espresse dal Cinquecento in poi si manifestino anche indicatori adatti a intraprendere una descrizione seria e scientifica (cfr. *infra*).
Quali erano le opinioni – stereotipe o scientificamente approfondite – sulle lingue europee? Alla ricerca di una risposta a una tale domanda ci si rende conto che si ha a che vedere con tutto un intreccio di fattori (v. fig. 2) che potevano aver causato un tale parere dotto. Comprendono correnti culturali e storiche internazionali, stereotipi nazionali, vicende nell'ambito della storia della musica, della politica o della religione, ma anche l'evoluzione e lo status delle lingue europee, se pensiamo alla concorrenza tra francese e italiano.

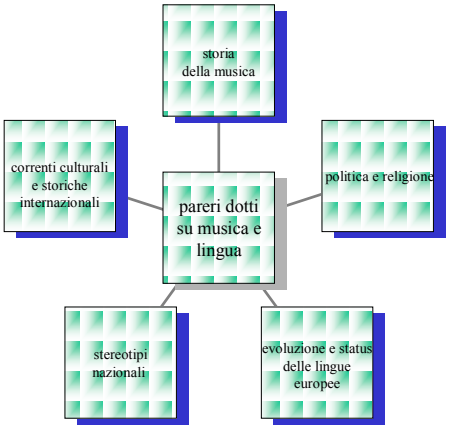


Figura 2: Contesti storico-culturali dei pareri dotti su musica e lingua

Consideriamo alcuni esempi espressi da varie persone illustri sull'estetica di alcune lingue (parlate). Il primo ad aver introdotto le formule stereotipiche sul suono delle lingue sembra essere stato un discepolo di Erasmus, il teorico tedesco Andreas Ornithoparcus, che nel 1517 pubblicò il suo trattato musicologico, *Musicae acivae micrologus libris quator digestis*, tradotto in inglese nel 1609. In questa opera troviamo per la prima volta il motivo di "Gallus cantat" (i francesi cantano quando parlano). L'elenco seguente dà un'idea della popolarità di questo topos, allargato successivamente ad altre lingue ed altri popoli[1]:

1517 Andreas Ornithoparchus (Germania):
„das Jauchzen der Engländer, den feinen Gesang der Franzosen, das Jammern der Spanier, das Meckern und Bellen der Italiener, das Heulen der Deutschen" (zitiert nach Biehle, 1931: 24);
1556 Hermann Finck (Germania):

[1] Le fonti seguenti sono tutte elencate nella bibliografia del nostro progetto di ricerca "Lingua e Musica" dell'università di Düsseldorf. La bibliografia fa parte della monografia sullo stesso tema (in preparazione).

„Germani boant: Itali balant: Hispani eiulant: Galli cantant“;
1579 Henri Estienne (Francia):
„Balant Itali, gemunt Hispani, ululant Germani, cantant Galli“;
1670 Saint-Evremond (Inghilterra):
„Hispanus flet, dolet Italus, Germanus boat, Flander ululat, solus Gallus cantat“;
1671 Claude-Francois Ménestrier (Francia):
„Galli cantant, Hispani latrant, Itali caprizant, Germani boant“ (que les Francois chantoient, que les Espagnols abboyoient ou glapissoient, que les Italiens chevrottoient, & que les Allemans meugloient);
1704 Lecerf de la Vieville (Francia):
„Hispanus flet, dolet Italus, Germanus boat, Flander ululat, & solus Gallus cantat“;
1705 Franois Raguenet (Francia):
L‘Espagnol pleure, l‘italien se plaint, l‘allemand beugle, le Flammand hurle, & le seul Francois chante“;
1739 Johann Mattheson (Germania):
„Germani boant, Itali balant, Hispani eiulant, Galli cantant“ („die Teutschen blöcken, die Welschen blecken, die Spanier heulen, nur die Franzosen allein singen“);
1773 Christoph Willibald Gluck (Germania):
„C’est la raison pour laquelle je n’emploie point *les trilles*, *les passages* ni *les cadences* que prodiguent les Italiens. Leur langue, qui s’y prête avec facilité, n’a donc à cet égard aucun avantage pour moi ; elle en a sans doute beaucoup d’autres : mais, né en Allemagne, quelqu’étude que j’aie pu faire de la langue italienne, ainsi que de la langue françoise, je ne crois pas qu’il me soit permis d’apprécier les nuances délicates qui peuvent faire donner la préférence à l’une des deux, & je pense que tout étranger doit s’abstenir de juger entr’elles; “
1852 Adrien le Roux de Lincy (Francia):
„Les Italiens pleurent, les Allemands crient et les Francois chantent“.

Come possono essere nati pareri di questo genere? I dibattiti sull’eccellenza e sul pregio dell’una o dell’altra lingua – parlata, bisogna sempre specificare, perché sono coinvolte ovviamente soprattutto caratteristiche prosodiche – possono davvero essere gli unici motivi per questi stereotipi? È vero che interferiscono, come raffigurato in tavola 2, anche fattori politici. A partire dal Settecento però, quando la *querelle des bouffons*, cioè la lite erudita sulla supremazia musicale del carattere dell’opera musicale – la tradizionale opera seria (francese) o la nuova opera buffa (italiana), si acutizzava sempre di più, spuntava un numero crescente di trattati appassionati e invettive ardenti su chi aveva la lingua più adatta per la musica (v. la famosa corrispondenza tra Deodati de Tovazzi e Voltaire nel 1761, ma anche il saggio di Rousseau, *Lettre sur la musique française*, uscito

nel 1753). Anche compositori della stessa epoca avevano qualche voce in capitolo, come possiamo vedere nei seguenti esempi:

„[...] il n'y a ni mesure ni mélodie dans la Musique Françoise [...] le chant François n'est qu'un aboyement continuel, insupportable à toute oreille non prévenue" (Rousseau, 1753);

„wenn nur keine französin italienische arien singete, ich würde ihr ihre französische blerrerrey noch verzeyhen, aber gute Musick zu verderben! – das ist nicht auszustehen" (L. Mozart, 1. 2. 1764);

„wenn nur die verfluchte französische sprache nicht so hundsfüttisch zur Musique wäre! – das ist was Elendes – die Teütsche ist noch göttlich dagegen – und dann erst die sänger und sängerinnen [...] sie singen nicht, sondern sie schreyen – heülen – und zwar aus vollem halse, aus der nase und gurgel" (W. A. Mozart, 9. 7. 1778);

„It is universally allowed that the Italian tongue is more sonorous, more sweet, and of more easy utterance, than any other modern language; and that the Music of Italy, particularly the vocal, perhaps for this reason, has been more successfully cultivated than any other in Europe" (Burney, 1789).

Si tratta dunque solo di atteggiamenti sciovinisti e di brutte esperienze individuali? O c'è una sostanza interpretabile non solo musicologicamente, ma anche linguisticamente in tutte queste citazioni?

3 Analisi linguistico-tipologica

Tenendo conto dei numerosi trattati sulla tematica di "lingua e musica" nel Settecento – nella banca dati del progetto di ricerca dell'università di Düsseldorf figurano gli spogli di 60 testi solo per questo secolo – presentiamo qui, in rappresentanza degli altri, l'essenza di uno dei più famosi trattati, quello di Rousseau, dal quale abbiamo estratto gli argomenti, adottati dall'autore per corroborare le sue tesi, che ci servono per individuarne il nesso con il lato

linguistico-tipologico (è ovvio che bisogna "tradurli" nella terminologia fonetica di oggi[2]).
Caratteristiche della cantabilità di una lingua secondo Rousseau (1753) (corsivo nostro):

Or s'il y a en Europe une langue propre à la musique, c'est certainement l'Italienne; car cette langue est *douce***,** *sonore***,** *harmonieuse***,** et *accentuée* plus qu'aucune autre, et ces quatre qualités sont précisément les plus convenables au chant:
Elle est *douce*, parce que [...] la *rencontre des consonnes y est rare et sans rudesse*, et qu'un très grand nombre de *syllabes n'y étant formées que de voyelles*, les *fréquentes élisions* en rendent la *prononciation plus coulante*:
elle est *sonore*, parce que la plupart des *voyelles* y sont *éclatantes*,
qu'elle n'a pas de *diphtongues composées*,
qu'elle a peu ou point de *voyelles nasales*,
et que les *articulations rares et faciles distinguent mieux le son des syllabes*, qui en devient plus net et plus plein. [...]".

(J. J. Rousseau, *Lettre sur la Musique Française*, 1753)

D'altra parte Rousseau specifica in più le caratteristiche che possano essere i motivi per la non cantabilità di una lingua (corsivo nostro):

„....une qui ne serait composée que de *sons mixtes*,
de *syllabes muettes, sourdes ou nasales*,
peu de voyelles sonores,
beaucoup de consonnes [...]"

(Rousseau, *Lettre sur la Musique Française*, 1753).

Senza poter giudicare in questa sede la giustezza di ciascuno degli argomenti, risulta, dalla caratterizzazione di Rousseau, l'importanza

[2] Utilizziamo come opere di riferimento (fra l'altro): Abercrombie (1967), Albano Leoni and Maturi (2002), Albano Leoni (2009), Bertinetto (1981, 1989), Canepari (2006), Chierchia (1986), Daniele and Patel (2003), Dauer (1983), De Dominicis (1999), Dufter (2004), Frank (1995), Hansen (1994), Herry (2001), Hess (1975), Hirst and Di Cristo (1998), Hurch (1996), Koesters Gensini (2007), Floccia *et al* (2003), Krenn (1987), Lichem (1969), Lonnemann (2006), Maas (1999), Maddieson (1984), Mairano and Romano (2007), Marotta (1985), Maturi (2006), Mayerthaler (1982), Meinhold and Stock (1980), Meisenburg and Selig (1998), Nespor (1994), Nespor and Vogel (1986), Pensom (1998), Philipp (1974), Pike (1945), Pompino-Marschall (1990), Ramus, Nespor and Mehler (1999), Redard (1989), Schafroth (1999), Scheer (1999), Schmid (1999), Terens (1998), Vennemann (1982), Verluyten (1984, 1988), Voghera and Mancini (1994).

del repertorio delle vocali, della loro distribuzione, della loro frequenza sintagmatica e della loro qualità, inoltre la possibilità di omettere vocali facoltative. E ne deriva anche la problematica, per la realizzazione musicale, di vocali non pronunciate e deboli (come lo *schwa*) e di cluster consonantici. Il secondo tipo di analisi mira quindi a spogliare la sostanza fonetica (qualità, quantità), i procedimenti fonetici-fonologici (fenomeni "sandhi", come elisione, apocope ecc.), la struttura sillabica e le regolarità fono(sin)tattiche, la prosodia (accento, unità ritmiche, effetti dell'assenza di intonazione) – con l'obiettivo di concepire una *griglia generica di criteri fonetici e prosodici* che permetta di descrivere ogni lingua secondo la sua qualità di essere cantata oppure di essere "cantabile", cioè di rivelare "flessibilità" di adattarsi a un qualsiasi tipo di ritmo e a una qualsiasi serie di note.
Questa griglia conterrà i tratti seguenti:

repertorio (qualità e quantità) e distribuzione delle vocali
quantità consonantica
struttura sillabica (tipi di sillabe, presenza/assenza di vocali deboli e di cluster consonantici, percentuale vocalica nelle sillabe, tendenza alla ritmicità sillabica)
accentuazione e unità ritmiche (schemi accentuali, struttura dei piedi più frequenti, tendenza alla ritmicità accentuale)
variabilità: lunghezza vocalica, sinalefe, dieresi e sineresi, sincope e apocope, prostesi, epentesi o epitesi vocalica o consonantica.

Ciascuno dei criteri può contribuire a una determinata qualità di cantabilità, che può significare in lingua A una certa rigidità ritmica e in lingua B la possibilità ridotta di apocopare vocali.
Una volta compilata la griglia dei fenomeni fonetici e fonologici bisogna applicarla a ciascuna delle lingue analizzate – nel nostro progetto in una prima fase alle lingue "classiche" dell'opera musicale, cioè l'italiano, il francese e il tedesco, in una seconda fase anche ad altre lingue – esemplificate con una scelta di arie di opere oppure, per utilizzare il termine tecnico, di "numeri musicali".

4 Analisi linguistico-musicologica

Per rendere ciascuna delle tre lingue quella originale e una delle lingue tradotte (o adattate alla musica originale) era necessario trovare opere liriche di cui la versione originaria era in italiano e rispettivamente in tedesco e francese. Cercando di scegliere opere della stessa epoca abbiamo definito un *corpus di base* che consiste nelle tre opere seguenti[3]:

4.1 Le opere analizzate

Mozart, *Le nozze di Figaro* (1786) (originale: italiano, versioni tradotte/trasposte in francese e in tedesco);
Gluck, *Iphigénie en Tauride* (1779) (originale: francese, versione tradotta/trasposta in tedesco, nel 1781, da Gluck stesso, e versione italiana (1957));
Mozart, *Die Zauberflöte* (1791) (originale: tedesco, versioni tradotte/trasposte in francese e in italiano);
Abbiamo definito così un corpus di base di nove versioni sulla base di tre opere diverse. Per aver un corpus complementare di francese – Gluck era tedesco, anche se compose *Iphigénie* su libretto di un autore francese, Nicolas-François Guillard – sarà analizzata un'altra opera francese, *Carmen* di Bizet (1875) (con le versioni tedesca e italiana)[4]. Il *corpus largo* comprende dunque 12 esecuzioni di quattro opere liriche.

4.2 Il metodo dell'analisi linguistico-musicologica

Va da sé che non è possibile esaminare le opere intere. Abbiamo rinunciato ai recitativi, anche se questi rappresentano casi particolari e

[3] Il genere musicale però non è uguale (*Nozze*: opera lirica/buffa; *Flauto magico*: *singspiel*; *Ifigenia in Tauride*: opera lirica/tragica).

[4] È vero che anche Mozart non era di madrelingua italiana, anche se compose *Le Nozze di Figaro* in italiano (su libretto di Lorenzo da Ponte), ma non risulta chiara la domanda fino a quale punto Gluck padroneggiasse il francese – anche se la lingua francese è la sua lingua preferita perché, così il compositore, «le poète me fournira le plus de moyens variés d'exprimer les passions: c'est l'avantage que j'ai cru trouver dans les paroles de l'opéra d'Iphigénie, dont la poésie m'a paru avoir toute l'énergie propre à m'inspirer de la bonne musique…» (Gluck, 1773).

degni di essere considerati a parte (v. Green, 2001; Caron and Morel, 2007), e prendiamo solo in considerazione i cosiddetti "numeri musicali", tra cui soprattutto arie, in alcuni casi anche cavatine e cori. Questi brani musicali vengono analizzati, ascoltandoli accuratamente, in relazione al testo originale del libretto o alla sua traduzione. Viene poi registrata in una banca dati ogni differenza tra le singole realizzazioni canore del testo e una registrazione ritenuta di media qualità (cioè di media velocità, di un registro neutro)[5], ottenuta facendo semplicemente leggere il testo. Si accumulano così, battuta per battuta, nota per nota, divergenze o convergenze tra lingua "parlata" e lingua cantata. La lingua che rappresenta il più gran numero di distacchi (tra cantato e parlato) sembra quella meno cantabile. La banca dati ci permette, tramite particolari modi di ricerca, di ricevere risultati complessivi:

per ogni tratto linguistico (*item*), per esempio il caso di una epitesi vocalica o il caso dell'accento inusuale;
per ogni numero musicale (arie, cavatine ecc.);
per ogni opera (differenziabile inoltre in opera originale e opera tradotta);
per ogni compositore.

Le deduzioni che potremo trarre dai nostri risultati saranno molteplici e vanno dai dati validi per una singola opera fino a conclusioni applicabili a una lingua[6].

4.3 Due esempi

4.3.1 Segno di "flessibilità": epentesi e epitesi di una vocale

Guardiamo due esempi (cfr. anche Schafroth, 2002): un caso di variabilità prosodica – qui un'epentesi o un'epitesi – è la "rivitalizzazione" di una vocale (e quindi di una sillaba intera) che

[5] Se il progetto di Caron and Morel (2007) avesse già dei risultati concreti e affidabili si potrebbe perfino pensare a basarsi sulla maniera di recitare del Settecento.

[6] Ci saranno da discutere anche altri risultati, come ad esempio il fatto che Gluck, nella sua versione (originale) francese dell'*Ifigenia*, ha composto un solo brano nel tempo di tre quarti, sapendo (probabilmente) che (a causa dell'ossitonia) è tra le misure ritmiche più difficili da cantare.

nella lingua parlata non è realizzata (perché è diventata muta) o solo in condizioni nettamente formali (come la recitazione di un'opera letteraria).
Prendiamo il caso dello *schwa* francese, che in tanti lessemi è sia facoltativo (fr. *p(e)tit*) sia virtuale (fr. *comte*), cioè realizzato solo in alcune condizioni diafasiche. Possiamo illustrare questo fenomeno con le battute 3 e 4 della "Cavatina" nelle *Nozze di Figaro* di Mozart: "mon pE-tit com-tE" (nell'originale italiano "si-gnor Con-ti-no").

Se	vuol	bal-	|	**la-**	re	|
Si	tu	veux	|	**dan-**	ser	|
Will	der	Herr	|	**Graf**	ein	|
1				2		
si-	gnor	Con-	|	**ti-**	no,	|
mon	pE-	tit	|	**com-**	tE,	|
Tänz-	chen	nun	|	**wa-**	gen,	|
	3				4	
il	chi-	tar-	|	**ri-**	no	|
de	la	gui-	|	**ta-**	rE	|
Mag	er's	mir	|	**sa-**	gen,	|
	13				14	
le	suo-	ne-	|	**rò**	sì,	|
je	te	joue-	|	**rai**	oui,	|
ich	spiel	ihm	|	**auf**,	ja,	|
	15				16	

Figura 3: Mozart, *Le nozze di Figaro*: "Cavatina", battute 1-4 e13-16

È solo grazie a questa "flessibilità" del francese che le parole possono riempire tutta la misura del tempo di tre quarti.

4.3.2 Segno di “non flessibilità”: l’accento fisso

Un altro caso costituisce l’accento: quando in un’aria d’opera l’accento principale cade sulla prima unità ritmica (semiminima o, linguisticamente, sillaba), nel tempo di tre quarti per esempio (che corrisponde al dattilo), una lingua come l’italiano, grazie al suo accento libero, offre molte più soluzioni lessicali al librettista rispetto al francese, che è caratterizzato da un accento quasi fisso, cioè ossitono. Nella prima battuta la congiunzione *se* accentuata pienamente può parere, rispetto a una realizzazione fonica media (dove ci sarebbe al massimo un accento secondario), leggermente innaturale, mentre l’accento parossitono in *ballare* corrisponde perfettamente alle regole prosodiche dell’italiano. Anche nel tedesco parlato l’ausiliare *will* potrebbe almeno portare un accento secondario, invece nella realizzazione cantata ha un accento principale; in più la parola *Graf* è accentuata sia nella frase letta sia qui nella cavatina: quindi in questo esempio c’è una notevole corrispondenza tra parlato e cantato in tedesco. In francese sia la congiunzione *si* sia il verbo *danser* appaiono prosodicamente marcati rispetto ad una realizzazione fonica. E c’è da chiedersi se Mozart si riferisse a quella sensazione uditiva quando criticava violentemente la lingua francese.

Riferimenti bibliografici

Discografia

Jourdan, P. (1997) *W. A. Mozart. Les Noces de Figaro*. Nouvelle adaptation en français d’Eric-Emmanuel Schmitt, en collaboration avec Pierre Jourdan (Orchestre de Chambre de l’Opéra d’État Hongrois Failoni; Ensemble Mille e Tre; Théâtre Français de la Musique – Théâtre Impérial de Compiègne). 3 CD, Éditions TFM.

Marriner, N. (1985) *Wolfgang Amadeus Mozart, Le Nozze di Figaro* (José van Dam, Barbara Hendricks, Ruggero Raimondi, Lucia Popp; Ambrosian Opera Chorus; Academy of St. Martin-in-the-Fields). Excerpts, 1 CD, Philips (Digital Classics) 1986.

Suitner, O. (1964) *Wolfgang Amadeus Mozart. Die Hochzeit des Figaro*. Querschnitt in deutscher Sprache. Übersetzung von Hermann Levi (Walter Berry, Anneliese Rothenberger, Hilde

Güden, Hermann Prey; Chor der Staatsoper Dresden; Staatskapelle Dresden). 1 CD, Berlin Classics 1995.

Libretti

Mozart. Le nozze di Figaro/Die Hochzeit des Figaro. Opera buffa in vier Akten. Text von Lorenzo da Ponte. KV 492. Deutsche Übersetzung von Kurt Honolka. Klavierauszug, nach dem Urtext der Neuen Mozart-Ausgabe. Kassel u.a.: Bärenreiter 1976.

Mozart. Les Noces de Figaro ("Livret intégral", 24-139), Paris 1999. [= L'Avant-Scène Opéra n° 135-136].

Letteratura secondaria

Abercrombie, D. (1967) *Elements of General Phonetics*. Edinburgh: Edinburgh University Press.

Albano Leoni, F. (2009) *Dei suoni e dei sensi. Il volto fonico delle parole*. Bologna: Il Mulino.

Albano Leoni, F. and P. Maturi (2002) *Manuale di fonetica*. Roma: Carocci.

Bertinetto, P. M. (1981) *Strutture prosodiche dell'italiano*. *Accento, quantità, sillaba, giuntura, fondamenti metrici*. Firenze: Accademia della Crusca.

– (1989) 'Reflections on the dichotomy 'stress' vs. 'syllable-timing''. *Revue de Phonétique Appliquée* 91/93, pp. 99-130.

Biehle, H. (1931) *Die Stimmkunst.* Vol. 1: *Geschichtliche Grundlagen*. Leipzig: Kistner&Siegel.

Burney, Ch. (1789) *A general history of music. From the earliest ages to the present period*, vol. 4. London: Payne and Son, Robson and Clark, G.G.J. and J. Robinson (4 vols., vol. 1: 1776).

Canepari, L. (2006) *Avviamento alla fonetica*. Torino: Einaudi.

Caron, Ph. and M. Morel (2007) 'Note récapitulatives sur le programme de reconstitution de la diction haute du français vers 1700', in *Annales de l'Association pour un Centre de Recherche sur les Arts du Spectacle aux XVIIe et XVIIIe siècles* 2.

Chierchia, G. (1986) 'Length, syllabification and the phonological cycle in Italian'. *Journal of Italian linguistics* 8, 1, pp. 5-33.

Daniele, J.R. and A.D. Patel (2003) 'An empirical comparison of rhythm in language and music'. *Cognition* 87, 1, B35-B45.

Dauer, R. M. (1983) 'Stress-timing and syllable-timing reanalyzed'. *Journal of Phonetics* 11, pp. 51-62.

De Dominicis, A. (1999[2]) *Fonologia comparata delle principali lingue europee moderne*. Bologna: CLUEB.

Deodati de Tovazzi, G.L. (1761) *Dissertation sur l'excellence de la langue italienne* [suivi par la *Lettre de M. de Voltaire à M. Deodati de Tovazzi, au sujet de la dissertation sur l'excellence de la langue italienne, avec la réponse de M. Deodati à M. De Voltaire*]. Paris: Bauche, Leclerc, Lambert.

Dufter, A. (2004) Ist das Französische eine silbenzählende Sprache?, in T. Meisenburg and M. Selig (eds) *Nouveaux départs en phonologie. Les conceptions sub- et suprasegmentales*, Tübingen: Narr, pp. 139-159.

Floccia, C., R. Kolinsky, C. Dodane and J. Morais (2003) 'Discriminating spoken words in French: the role of the syllable and the CV phonological skeleton'. *Language and cognitive processes* 18, pp. 241-67.

Frank, B. (1995) *Die Lautgestalt des Französischen. Typologische Untersuchungen*. Tübingen: Niemeyer.

Gluck, Ch. W. Ritter von (1773) 'Lettre de M. le chevalier Gluck, sur la Musique'. *Mercure de France* Février 1773, pp. 182-184. Genève: Slatkine Reprints 1971.

Green, E. (2001) *La parole baroque*. *Un essai*. Paris: Desclée de Brouwer.

Hansen, A.B. (1994) 'Etude du E caduc – stabilisation en cours et variations lexicales'. *French Language Studies* 4, pp. 25-54.

Herry, N. (2001) *Evaluation subjective et objective de la prosodie anglaise parlée par des français: Apport de l'enseignement assisté par ordinateur*. Thèse de doctorat, Université de Provence.

Hess, Z. (1975) *Typologischer Vergleich der romanischen Sprachen auf phonologischer Basis*. Frankfurt am Main-Bern: Lang.

Hirst, D. and A. Di Cristo (eds) (1998) *Intonation Systems. A Survey of Twenty Languages*. Cambridge: Cambridge University Press.

Hurch, B. (1996) Accentuations, in B. Hurch and R.A. Rhodes (eds) *Natural Phonology: The State of The Art*, Berlin-New York: Mouton de Gruyter, pp. 73-96.

Koesters Gensini, S.E. (2007) *Fonetica e fonologia del tedesco*. Bari: B.A. Graphis.

Krenn, H. (1987) '*... Del bel paese là dove il sì suona...* Zur Originalität italienischer Lautung im Vergleich mit anderen europäischen Sprachen'. *Italienische Studien* 10, pp. 121-32.

Lichem, K. (1969) *Phonetik und Phonologie des heutigen Italienisch*. München: Hueber.

Lonnemann, B. (2006) *Schwa, Phrase und Akzentuierung im français du Midi – eine kontrastive Untersuchung im Rahmen des Projektes „La Phonologie du français contemporain (PFC): usages, variétés et structure“.* Osnabrück. Available from http://elib.ub.uni-osnabrueck.de/publications/diss/E-Diss606_thesis.pdf (accessed 20 September 2009).

Maas, U. (1999) *Phonologie*. *Einführung in die funktionale Phonetik des Deutschen*. Opladen: Westdeutscher Verlag.

Maddieson, I. (1984) *Patterns of sounds*. Cambridge: Cambridge University Press.

Mairano, P. and A. Romano (2007) Inter-subject agreement in rhythm evaluation for four languages (English, French, German, Italian), in J. Trouvain and W.J. Barry (eds) *Proceedigs of the 16th International Congress of Phonetic Sciences* (Saarbrücken, Germania, 6-10 agosto 2007), Saarbrücken: Univ. des Saarlandes (CD-ROM), pp. 1149-52.

Marotta, G. (1985) *Modelli e misure ritmiche: la durata vocalica in italiano*. Bologna: Zanichelli.

Maturi, P. (2006) *I suoni delle lingue, I suoni dell'italiano. Introduzione alla fonetica*. Bologna: Il Mulino.

Mayerthaler, E. (1982) *Unbetonter Vokalismus und Silbenstruktur im Romanischen*. *Beiträge zu einer dynamischen Prozeßtypologie*. Tübingen: Niemeyer.

Meinhold, G. and E. Stock (1980) *Phonologie der deutschen Gegenwartssprache*. Leipzig: Bibliographisches Institut.

Meisenburg, T. and M. Selig (1998) *Phonetik und Phonologie des Französischen.* Stuttgart: Klett.

Mozart. Briefe und Aufzeichnungen. Gesamtausgabe, 7 Bde, hg. von der Internationalen Stiftung Mozarteum Salzburg, gesammelt und erläutert von W.A. Bauer, O.E. Deutsch & J. Heinz, Kassel: Bärenreiter 1962-75.

Nespor, M. (1994) *Fonologia*. Bologna: Il Mulino.

Nespor, M. and I. Vogel (1986) *Prosodic Phonology*. Dordrecht: Foris.

Pensom, R. (1998) *Accent and metre in French. A theory of the relation between linguistic accent and metrical practice in French, 1100-1900*. Bern: Lang.

Philipp, M. (1974) *Phonologie des Deutschen*. Stuttgart: Kohlhammer.
Pike, K.L. (1945) *The Intonation of American English*. Ann Arbor: University of Michigan Press.
Pompino-Marschall, B. (1990) *Die Silbenprosodie. Ein elementarer Aspekt der Wahrnehmung von Sprachrhythmus und Sprechtempo*. Tübingen: Niemeyer.
Ramus, F., M. Nespor and J. Mehler (1999) 'Correlates of linguistic rhythm in the speech signal'. *Cognition* 73, 3, pp. 265-92.
Redard, F. (1989) 'Comparaison des systèmes phonologiques de l'italien et du français'. *Studi italiani di linguistica teorica e applicata* 18, 1-2, pp. 115-117.
Rousseau, J.-J. (1753) Lettre sur la musique française, in B. Gagnebin and M. Raymond (eds) *Jean-Jacques Rousseau: Œuvres complètes*, vol. 5: *Écrits sur la musique, la langue et le théâtre*, Paris: Gallimard 1995.
Schafroth, E. (1999) 'Zur Virtualität des [ə]. Forschungsbericht, Fragen und Ergebnisse zu einem bekannten Phänomen'. *Zeitschrift für französische Sprache und Literatur* 109, pp. 113-147.
– (2002) Sprache und Musik. Sprachwissenschaftliche Beobachtungen zur Opera buffa *Le nozze di Figaro* und ihren deutschen und französischen Fassungen, in S. Heinemann, G. Bernhard and D. Kattenbusch (eds) *Roma et Romania. Festschrift für Gerhard Ernst zum 65. Geburtstag*, Tübingen: Niemeyer, pp. 287-304.
Scheer, T. (1999) 'Aspects de l'alternance schwa-zéro à la lumière de 'CVCV''. *Recherches linguistiques de Vincennes* 28, pp. 87-114.
Schmid, S. (1999) *Fonetica e fonologia dell'italiano*. Torino: Paravia.
Ternes, E. (1998) Lauttypologie der Sprachen Europas, in W. Boeder, C. Schroeder, K. Heinz Wagner und W. Wildgen (eds) *Beiträge zur empirischen Sprachwissenschaft*, Tübingen: Narr, pp. 139-52.
Vennemann, Th. (ed) (1982) *Silben, Segmente, Akzente*. Tübingen: Niemeyer.
Verluyten, S.P. (1984) Phonetic Reality of Linguistic Structures: the Case of (Secondary) Stress in French, in M.P.R. Van den Broecke and A. Cohen (eds) *Proceedings of the Tenth International Congress of Phonetic Sciences*, Dordrecht: Foris, pp. 522-6.
Verluyten, S.P. (ed.) (1988) *La phonologie du schwa français*. Amsterdam-Philadelphia: John Benjamins.

Voghera, M. and F. Mancini (1994) ‘Lunghezza, tipi di sillabe e accento in italiano’. *Archivio glottologico italiano* 79, 1, pp. 51-77.
Voltaire (1761) Lettre à G. L. Deodati de Tovazzi du 24 janvier 1761, in Th. Besterman (ed.) (1964), *Voltaire: Correspondance*, vol. VI (octobre 1760-décembre 1762), Paris: Gallimard, pp. 231-35.

Sistemi fonetici in contatto

La variabilità di /r/ nell'italiano di tedescofoni altoatesini

Lorenzo Spreafico, Alessandro Vietti
Libera Università di Bolzano
lorenzo.spreafico@unibz.it; alessandro.vietti@unibz.it

1 Introduzione[1]

Gli obiettivi dello studio che verrà qui presentato sono (a) di tracciare un quadro preliminare dei rotacismi nell'italiano parlato in Alto Adige e (b) di analizzare in particolare il sottoinsieme di varianti presenti nell'italiano di due parlanti tedescofoni. La descrizione e l'analisi che seguono costituiscono pertanto un primo passo nella direzione di definire con maggiore chiarezza e sulla base di solidi riscontri empirici i complessi rapporti linguistici e sociolinguistici tra le varietà di italiano e tedesco presenti nel Tirolo meridionale[2]. Se si prende in esame, per esempio, l'italiano parlato in Alto Adige si possono immediatamente distinguere diverse varietà di lingua riconducibili a distinte classi di utenti e contraddistinte da tratti linguistici specifici. Tracciando un quadro sommario, nei contesti urbani altoatesini si possono individuare:
a) la varietà regionale di italiano, con tracce più o meno evidenti dell'influsso dei dialetti veneti meridionali (STI-i);
b) un *continuum* di varietà di italiano di tedescofoni (da quelle intermedie fino alle *native-like*) nelle quali traspare, a vari livelli, l'influenza del sostrato dialettale tirolese (STI-d);
c) la varietà di italiano parlato da parlanti bilingui, delle quali invece si sa ancora molto poco (STI-b).
Lo scopo di questo articolo, pur collocandosi all'interno di un'indagine più vasta sull'italiano in Alto Adige, è circoscritto allo

[1] Sebbene il lavoro sia il risultato di riflessioni comuni, la stesura dei paragrafi 1, 4.2 e 4.3 è stata curata da Alessandro Vietti, mentre quella dei paragrafi 2, 3, 4.1 da Lorenzo Spreafico.

[2] La presente indagine fa parte di un progetto pilota di descrizione fonetica strumentale dei rotacismi in Alto Adige finanziato con fondi d'Ateneo per gli anni 2009 e 2010 (BW5033).

studio dei foni /r/ nelle varietà di tedescofoni altoatesini. Dopo aver definito il campo di variazione dei suoni /r/ in Alto Adige, si procederà quindi all'analisi delle realizzazioni uvulari (/R/) tipiche dell'italiano di tedescofoni (d'ora in avanti STI-d) ponendo in relazione la variabilità degli allofoni con il contesto distribuzionale (nella catena fonica e all'interno nella sillaba), con il parlante e con il compito comunicativo (lettura di parole e frasi vs map task).

2 /r/ in Italia e in Alto Adige

La classe di suoni /r/ tende a essere trattata unitariamente, talvolta per via della possibilità di individuare tra i suoi membri caratteristiche fonologiche comuni (Wiese, 2001), ma più spesso in conseguenza della propensione nei sistemi ortografici di derivazione latina a impiegare un medesimo carattere per trascriverla (Ladefoged and Maddieson, 1996: 215). Ciononostante tale classe ricomprende al suo interno foni estremamente diversificati sia dal punto di vista articolatorio - si pensi anche solo alla possibilità di identificare svariati luoghi (labiodentale, alveolare, postalveolare, velare, uvulare) e modi (vibrante, fricativo, approssimante, vocalico) di articolazione -; che dal punto di vista acustico (cfr. a tal proposito tra gli altri Recasens, 1991; Ladefoged and Maddieson, 1996; Docherty and Foulkes, 2001), tanto da rendere di fatto impossibile il riconoscimento di proprietà fonetiche comuni tra i suoi membri. Proprio la volontà di meglio comprendere le potenzialità e le motivazioni di tale variabilità ha recentemente spinto a indagare, con sempre maggiore attenzione e da più prospettive teoriche, le dimensioni di variazione di /r/, soprattutto in lingue indoeuropee quali il francese (Demolin, 2001; Romano, in press), l'inglese (Foulkes and Docherty, 2001; Stuart-Smith, 2007), il nederlandese (Van de Velde and Van Hout, 1999), il tedesco (Simpson, 1998: 105-150; Wiese, 2001), lo svedese (Engstrand, Frid and Lindblom, 2007).

Per quanto concerne l'italiano le descrizioni sono ancora piuttosto limitate e, salvo alcune eccezioni (Romano, in press; Vietti, Spreafico and Romano, in press; Vietti and Spreafico, 2008; Felloni, 2008), limitate a brevi paragrafi in descrizioni generali. In queste ultime si è solitamente concordi nell'affermare che la realizzazione non marcata di /r/ per l'italiano standard è costituita dalla polivibrante alveolare [r] (cfr., tra gli altri, Nespor, 1993; Schmidt, 1999; Maturi, 2006),

articolata facendo battere più volte l'apice della lingua contro gli alveoli. Talvolta nel trattare di [r] gli stessi autori fanno riferimento a possibili differenze di durata legate alla distinzione di lunghezza fonologica e, quindi, discutono la possibilità di contrapporre /r/ e /rr/ sulla base del numero di chiusure riscontrabili nell'articolazione del fono. Questo genere di osservazioni si rinviene sia in ricerche di marca strumentale che in indagini di impostazione uditiva. Per quanto concerne le prime rilevanti sono le osservazioni in Ladefoged and Maddieson (1996: 219-221) dove si riporta che, per quanto riguarda l'opposizione tra le due possibili realizzazioni in posizione intervocalica, l'analisi di un campione minimo rivela come la vibrante geminata sia resa tramite un numero di occlusioni compreso tra tre e sette, mentre la vibrante singola sia data da un numero di battiti mai superiore a due. Maturi (2006: 78), dal canto suo, computa per l'italiano rispettivamente 2-4 chiusure nel caso di /r/ e 5-7 in quello di /rr/. Canepari (1999: 98; 2006: 69) distingue invece sulla base dell'ascolto, tra due o – nella sola occasione di realizzazioni enfatiche – tre occlusioni. Tuttavia nelle medesime sedi l'autore discute anche della possibilità di articolare un fono *vibrato* [ɾ] caratterizzato da un solo battito che, nella pronuncia neutra dell'italiano, si troverebbe in distribuzione complementare con l'allofono [r]: mentre quest'ultimo occorrerebbe tendenzialmente in sillabe accentate del tipo (C/V)ˊrV-, ˊCrV,ˊVr:C(V),ˊV(.)r#, il monovibrante [ɾ] comparirebbe in quelle prive di accento primario comeˊV:ɾV, (V/C)ɾV, -Vɾ[3].

Canepari (1999: 98-102) registra inoltre la presenza in italiano di quattro realizzazioni uvulari (polivibrante [ʀ], fricativa sonora [ʁ], approssimante [ʁ̞] e *vibrata*, quest'ultima definita come «suono più debole» (1999: 98) oppure «monovibrante» (2006: 70)[4] rinvenibili tanto nel caso di pronunce caratterizzate da difetti ortoepici, quanto in

[3] L'analisi, elaborata sulla base di ascolti non strumentali, è stata verificata da Romano (in press) che ha operato un controllo spettroacustico degli esempi riportati da Canepari (1999: 328) a supporto delle proprie osservazioni verificando come in effetti [ɾ] compaia tra vocali in posizione non accentata oppure nella fase esplosiva nella realizzazione di /rr/.

[4] Si preferisce però rinunciare qui, così come nel resto del contributo, alla classificazione e alla trascrizione [can]IPA.

alcune pronunce regionali[5] tra cui quella dell'Alto Adige che qui interessa trattare.

Le indagini sulla fonetica e la fonologia dell'italiano dell'Alto Adige sono invero limitate. In questo senso le ricerche fondamentali risultano essere Mioni (1990; 2001); Canepari (1990); Tonelli (2002).

Nel lavoro di Mioni (1990) l'analisi è però limitata alla sola norma cittadina, al parlato di studenti delle scuole superiori di lingua italiana[6] e tedesca. Per quanto riguarda il campione italofono la variante prevalente il tap apicoalveolare; invece per quanto riguarda il campione tedescofono l'autore rileva che gli informatori considerati usano «tutti e categoricamente» un qualche tipo di /r/ uvulare (1990: 203), identificando in ciò il più tipico caso di influsso del sostrato tedesco, ancora più evidente se si considera che, in generale, «la fonologia dell'italiano in questi informanti risulta assai ben acquisita» ed «è in sostanza quella degli italofoni circostanti, senza una rilevante presenza di modelli esterni alla regione» (1990: 204). Anche Canepari (1990: 393; cfr. anche 2006: 171) – trattando però dell'Alto Adige nel suo complesso – registra la presenza di pronunce tendenzialmente uvulari (costrittive, vibranti o approssimanti), seppur occasionalmente accompagnate da rese alveo-uvulari (come nel caso della vibrante, cfr. [can]IPA [ɾ̠]). Tonelli (2002: 50) riporta invece che l'unica variante di /r/ individuata in un campione di italiano (presumibilmente di italofoni) di Bolzano da lei analizzato è la monovibrante alveolare [ɾ], talvolta realizzata con un maggior numero di trilli in conseguenza di pronunce marcate.

Vietti and Spreafico (2008) offrono invece un quadro differente. Infatti, muovendo da un *corpus* costruito analogamente a quanto riportato in §3, essi indagano strumentalmente la natura delle realizzazioni di /r/ nelle produzioni di parlanti altoatesini rilevando la compresenza di rese apicali e uvulari non soltanto in parlanti provenienti da aree diverse dell'Alto Adige, ma talvolta anche nelle

[5] Omettiamo qui per ragioni di spazio la discussione delle caratteristiche di altre pronunce regionali per le quali cfr., oltre a Canepari (1999: 101-102 e i numerosi passi nei capitoli 9-15; 2006: 169-74), soprattutto Romano (in press) che offre un'accurata indagine anche strumentale tra le altre in particolare di toscano e piemontese e Vietti, Spreafico and Romano (in press) che tratta della resa del nesso /ˊori/ della parola *frigorifero* in un campione di dati semispontanei.

[6] In tal caso si tratta di immigrati italiani di seconda generazione.

produzioni di uno stesso informante. Analizzando un campione di 435 occorrenze ricavate dalla lettura di liste di frasi e parole da parte di undici informanti provenienti da diverse aree dell'Alto Adige, Vietti and Spreafico (2008) osservano che nel complesso i foni consonantici utilizzati sono almeno quattordici ovvero, in ordine decrescente di frequenza: tap alveolare [ɾ]; tap uvulare [tap ʀ]; approssimante uvulare [ʁ̞]; trillo uvulare [ʀ]; fricativa uvulare sorda [χ]; polivibrante alveolare [r]; flap alveolare [flap ɾ]; flap retroflesso [ɽ]; fricativa retroflessa sonora [ʐ]; approssimante retroflessa [ɻ]; fricativa uvulare sonora [ʁ]. Accanto a questi si collocano poi numerosissimi casi di cancellazione, oltre che fenomeni di difficile classificazione dovuti soprattutto alla scarsa intelligibilità del dato uditivo e di quello strumentale che indurrebbero spesso – ma gli indizi in tal senso sono sempre ambigui o insufficienti – a ritenere trattarsi di articolazioni alveo-uvulari.

Varianti	Numero di parlanti
4	*2*
5	*4*
6	*2*
7	*2*
8	*1*

Tabella 1: Realizzazioni differenti per numero di parlanti (Vietti and Spreafico, 2008)

Vietti and Spreafico (2008) rilevano inoltre che ciascun parlante utilizza sempre almeno quattro foni differenti o che, addirittura, in taluni casi ne impiegano sette o otto (Tab. 1) e cercano di comprendere se tale variazione sia legata a fattori linguistici o extralinguistici. Operando un'analisi statistica multivariata[7] Vietti and Spreafico (2008) giungono a osservare tra l'altro che, per quanto concerne i primi fattori, i tap alveolari occorrono in ciascun contesto possibile mentre le approssimanti uvulari e le cancellazioni si presentano solo in coda di sillaba o in prossimità di confini di parola. Per quanto concerne invece i fattori extralinguistici gli autori invocano, soprattutto in relazione alla presenza di rese uvulari, l'esser

[7] Per maggiori informazioni sulla tecnica cfr. §4.2.

parlanti di una varietà di *STI-i*, *STI-d* o *STI-b* (cfr. §1), e quindi, indirettamente, l'azione dell'eventuale sostrato tedesco.
In tal senso va osservato che, purtroppo, pure in questo caso si registra una carenza di ricerche sistematiche, di fatto limitate alle osservazioni di Tonelli (2002: 42) sullo *Bozner Deutsch* – la varietà di dialetto di tedesco bavarese parlata nel capoluogo altoatesino caratterizzata da prevalenti rese vibranti uvulari e, solo marginalmente, alveolari – e soprattutto alle indagini dello *Tirolischer Sprachatlas* (Klein and Schmitt, 1969). Per quanto riguarda le realizzazioni di /r/ lo spoglio dei dati in esso contenuti permette di osservare una variazione diatopica rilevantissima e caratterizzata da considerevoli oscillazioni persino in seno alla stessa valle – si pensi alle cancellazioni o alle deboli realizzazioni apicali della Pusteria occidentale contrapposte alle marcate polivibranti uvulari di quella orientale[8] – di fatto in linea con quella già osservata per altre aree tedescofone (per una visione d'insieme cfr. Simpson, 1998: 105-15). Per quanto concerne le aree qui rilevanti perché di provenienza degli informanti di cui si riferirà sotto, l'osservazione delle carte dell'atlante permette di identificare una prevalente realizzazione (polivibrante) uvulare nel caso del Burgraviato (Merano) e (polivibrante) apicale in quello della Valle Isarco (Bressanone) e della già considerata Pusteria occidentale (Brunico).

3 Metodologia

I dati analizzati sono estratti da una raccolta di registrazioni[9] di italiano parlato in Alto Adige in corso di elaborazione presso la Libera Università di Bolzano (*FONIAA, Fonetica e Fonologia dell'Italiano in Alto Adige*). La banca dati attualmente comprende le produzioni di ventotto informanti di età compresa tra i 25 e i 42 anni riconducibili a tre distinti gruppi di parlanti: quello degli italofoni (n=7), quello dei tedescofoni (n=11) e, infine, quello dei bilingui italiano-tedesco (n=7). Le circa sette ore di registrazioni disponibili sono state elicitate

[8] È interessante osservare come nella Pusteria orientale si registrino anche fricative uvulari – caratteristiche del solo parlato di persone anziane – e rese velouvulari (Klein and Schmitt, 1969: 2. Band, Karte 50, 51, 53, 54).
[9] La raccolta è stata effettuata impiegando un registratore Marantz PMD660 e un microfono Behringer B1. La campionatura ha avuto luogo a 22 kHz, la digitalizzazione a 16 bit.

impiegando tecniche differenti, così da ottenere campioni di parlato dallo stile distinto: anzitutto quello dialogico semispontaneo – ottenuto chiedendo a coppie di informanti di risolvere un *map task* – quindi quello monologico, ricavato chiedendo loro di leggere una lista di parole e una lista di frasi. In entrambi i casi si è trattato di materiali originali, seppur elaborati muovendo da quelli impiegati nel progetto CLIPS, sviluppati per raccogliere quante più occorrenze possibile di /r/ nei diversi contesti consentiti in italiano.
I dati di cui nei prossimi paragrafi presentiamo l'analisi sono stati ripresi da quelli di due parlanti tedescofoni che presentano nelle loro produzioni unicamente realizzazioni uvulari (o cancellazioni) di /r/. JAS è un'informante di 29 anni, figlia di parlanti nativi di dialetto tirolese, nata e vissuta a Bressanone sino a 11 anni, quando si è trasferita con la famiglia a Brunico, dove ha frequentato le scuole medie inferiori e superiori di lingua tedesca. E' impiegata amministrativa a Bolzano.
CHR, figlio di parlanti nativi di dialetto tirolese, è un informante di 31 anni nato e cresciuto a Lagundo. Ha compiuto l'intero percorso scolastico, concluso con un diploma di scuola professionale, presso le scuole di lingua tedesca di Merano. Svolge mansioni tecniche presso un'amministrazione pubblica di Bolzano.

4 Analisi dei dati

4.1 Classificazione spettroacustica delle varianti

I dati raccolti sono stati classificati alla luce di un'analisi spettroacustica svolta servendosi di PRAAT[10] così da riportare a classi discrete di foni le diverse realizzazioni di /r/ riconosciute nel campione. In complesso sono state individuate 385 rese di /r/: 264 sono costituite da realizzazioni consonantiche; 61 da cancellazioni; 40 da rotacizzazioni vocaliche; 20 da foni (solitamente consonantici) difficilmente classificabili per via di difetti di registrazione oppure della scarsa leggibilità spettroacustica (Tab. 2).

[10] Versione 5.1.07.

	ʀ tap	ʀ	ʀ̥	ʁ	ʁ̞	ʁ̥	χ	Vˠ	cancellazione	indecidibile	totale
JAS	*3*	*2*	*1*	*9*	*55*	*1*	*17*	*32*	*56*	*7*	*183*
CHR	*51*	*45*	*3*	*6*	*18*	*/*	*52*	*8*	*5*	*13*	*201*

Tabella 2: Realizzazioni di /r/ in CHR e JAS

L'analisi qualitativa delle rese consonantiche permette di osservare che in nessun caso si registrano articolazioni alveolari del tipo più comune in italiano standard. Al contrario, le realizzazioni sono unicamente uvulari – proprio come riportato per le varietà di dialetto tirolese che costituiscono le L1 degli informanti – e prevedono l'adozione di quattro distinti modi di articolazione: vibrante, tap, fricativo, approssimante.

I foni più chiaramente riconoscibili sono i tap e i trilli uvulari, entrambi caratterizzati acusticamente dalla presenza di una (tap) o più (trilli) discontinuità nel tracciato spettrografico e da una o più cadute del valore dell'intensità dovute all'assenza o al repentino smorzamento del segnale in corrispondenza delle chiusure determinate dal contatto tra il palato molle e il postdorso della lingua. Molto spesso – ed è per esempio il caso riportato in Fig. 1 relativo alla lettura di <carro attrezzi>, *item* incluso nella lista di parole (Carro_AttRezzi, JAS) – lo spettrogramma conserva in coincidenza delle occlusioni tracce della struttura formantica ad indicare che la chiusura è solo parziale, probabilmente anche per via dell'esigua massa dell'ugola[11].

[11] Stante la natura eminentemente qualitativa dell'indagine non è stato rilevato alcun parametro caratterizzante per i tap né, soprattutto, per i trilli. Segnaliamo comunque che per quanto riguarda la parlante JAS le risonanze relative alla terza formante dei foni vibranti si collocano attorno a valori tendenzialmente superiori rispetto a quelli indicati da Lafoged and Maddieson (1996: 226) - ovvero in prossimità dei 3500Hz - e che, perlomeno nel parlato spontaneo, le chiusure non sono mai più di quattro.

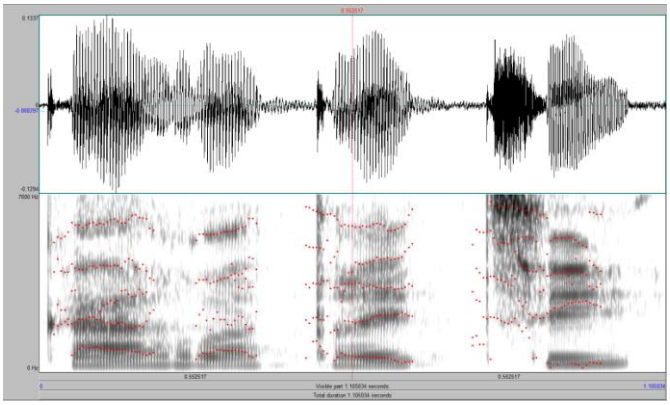


Figura 1: trillo e tap uvulare, <carro attrezzi> JAS

Anche le fricative uvulari (Fig. 2) presentano correlati acustici tipici che le rendono immediatamente distinguibili. Infatti sono caratterizzate – sia sul tracciato della forma d'onda che su quello spettrografico – da un segnale irregolare ma continuo indicatore delle turbolenze del flusso d'aria emesso. Data la rilevante estensione della cavità oltre il *locus* di costrizione, le turbolenze visualizzate risultano occupare bande di frequenza piuttosto basse, collocate rispettivamente per i due parlanti all'interno dei 3600-6100Hz (CHR) e dei 3100-5600Hz (JAS)[12]. Meno immediata è invece l'attribuzione del valore di sonorità del fono analizzato. Infatti, molto spesso, il caratteristico segnale periodico dovuto alla vibrazione delle corde vocali scompare ma non si accompagna ad alcun accrescimento nell'intensità del segnale, ovvero ad alcun indizio concorrente di assordamento del fono.

[12] Si tratta, anche in questo caso, di valori orientativi risultanti da osservazioni impressionistiche e non di medie esito di misurazioni sistematiche.

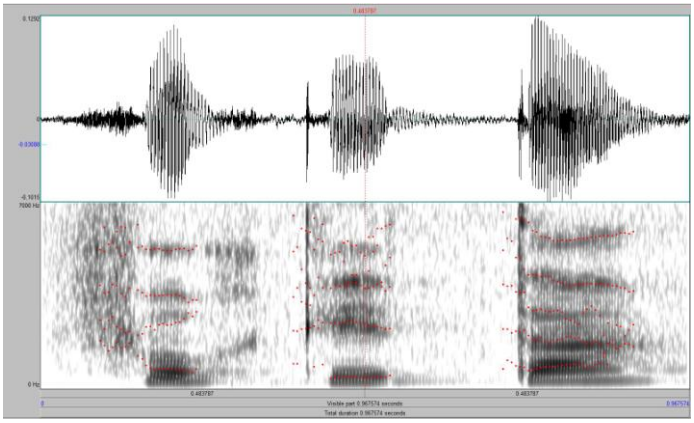


Figura 2: fricativa uvulare sorda, <forchetta> JAS

Le approssimanti uvulari (Fig. 3) rappresentano articolazioni di più difficile caratterizzazione, in particolare se rapportate alle rese con vocale rotacizzata. Il dato acustico dei due tipi di foni è infatti molto simile e distinto solo nei valori di intensità che, nel caso delle approssimanti, risultano notevolmente smorzati e comportano l'indebolimento – se non addirittura la cancellazione – delle tracce relative all'andamento di F3 e F4. La perdita di energia che caratterizza le approssimanti, tuttavia, non è mai tale da far dubitare che il fono sia da classificare come tap oppure fricativo. Infatti, mentre nel primo caso la caduta dell'intensità non è sufficientemente rapida e marcata; nel secondo, l'assenza di tracce di frizione fuga ogni dubbio.

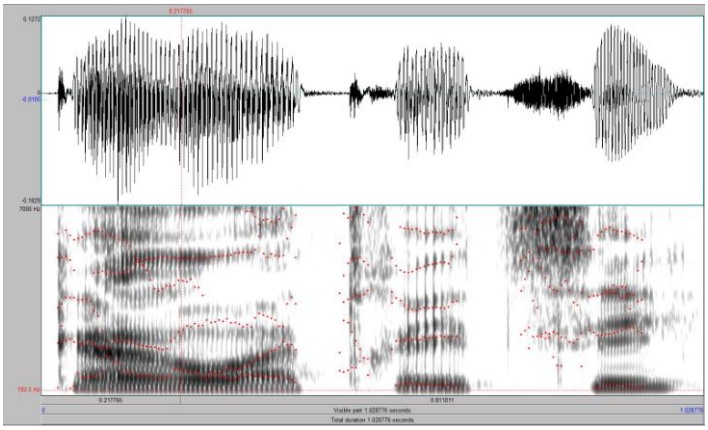


Figura 3: approssimante uvulare, <carro (attrezzi)> CHR

4.2 Variabili e metodo di analisi statistica

Su un corpus di 364[13] esecuzioni di /R/ è stata condotta un'analisi statistica multivariata con lo scopo primario di esplorare possibili associazioni tra le variabili in esame, ovvero tra i diversi allofoni di R (Varianti), la loro posizione nella struttura sillabica (Sillaba), la loro posizione nella catena fonica (Contesti fonetici) e il compito escussivo (Stile).
Come già osservato nella sezione precedente, la variabile Varianti è composta dai diversi tipi di foni rotici uvulari, classificati in categorie discrete sulla base dell'ispezione degli spettrogrammi. Un'analisi acustica unificata dei rotacismi non è ovviamente possibile, poiché la variazione allofonica che caratterizza questo "fonema" lo scompone in classi di foni con proprietà acustiche e articolatorie distinte, dalle fricative alle vibranti, da approssimanti a modulazioni vocaliche, fino alla cancellazione. Ciò detto, in realtà, pur all'interno di questo quadro apparentemente eterogeneo di suoni, l'analisi strumentale rivela sorprendentemente se non un vero e proprio *continuum* una sorta di *gradatum* con molte zone di transizione, caratterizzate da strategie articolatorie parallele e multiple. La labilità dei confini tra un tap e un'approssimante uvulari e tra quest'ultima e una fricativa sonora hanno reso arduo il compito di catalogazione dei foni in classi, ma hanno in un'ultima analisi rafforzato la congettura, già avanzata in Schiller (1998), circa la presunta unitarietà articolatoria di questo insieme di allofoni, determinata dalla maggiore o minore costrizione operata dal post-dorso della lingua sul palato molle, nella regione velare-uvulare.
Le categorie di foni uvulari analizzati sono sette: la fricativa sorda e sonora, la vibrante, il tap, l'approssimante, la rotacizzazione della vocale adiacente e la cancellazione (che a rigore non potrebbe essere definito come un fono R).
Le proprietà impiegate per definire gli allofoni sono, come già ricordato, di tipo distribuzionale nonché in parziale sovrapposizione. La prima, denominata per brevità Sillaba, riguarda la posizione di /R/ come testa o come coda all'interno dei macro-tipi sillabici ricorrenti

[13] Le realizzazioni classificate come indecidibili non sono state prese in considerazione.

nel corpus. Si sono pertanto distinti quattro tipi sillabici: in posizione di attacco sillabico RV, in testa preceduta da consonante CRV, in posizione di coda VR e, in ultimo, si è mantenuto come caso isolato quello nel quale nell'italiano standard si avrebbe un fono geminato VR.RV, anche se nell'area altoatesina non sono attestate /r/-/R/ geminate, sia nell'italiano regionale bolzanino (STI-i), che nella varietà di italiano di tedescofoni STI-d (v. Mioni 1990, Vietti and Spreafico 2008).
La variabile Contesti fonetici rappresenta invece l'intera gamma di contesti fonici attestati nel corpus ed è costituita da 75 valori, come si può osservare nella Tab. 3 che contiene i primi 14 contesti più frequenti (che rappresentano circa il 50% del corpus) .

	Frequenza	Freq. %	Freq. % cumulativa
tra	*26*	*6,8*	*6,8*
are	*20*	*5,2*	*11,9*
tre	*19*	*4,9*	*16,9*
ore	*18*	*4,7*	*21,6*
ora	*17*	*4,4*	*26,0*
aro	*16*	*4,2*	*30,1*
ero	*14*	*3,6*	*33,8*
tro	*11*	*2,9*	*36,6*
ori	*10*	*2,6*	*39,2*
ari	*9*	*2,3*	*41,6*
era	*9*	*2,3*	*43,9*
ere	*9*	*2,3*	*46,2*
ark	*8*	*2,1*	*48,3*
erk	*8*	*2,1*	*50,4*

Tabella 3: Contesti fonetici, i valori più frequenti

L'ultima variabile presa in considerazione è lo Stile o, più correttamente, il diverso compito comunicativo assegnato per l'elicitazione dei dati. I valori della variabile Stile, che corrispondono ai tre metodi di lettura di parole, lettura di frasi e *map task*, si pongono in questa fase esplorativa a un livello di generalità molto alto e nascondono al proprio interno fattori e indicatori eterogenei non soltanto per il grado di distanza dall'oggetto misurato, ma anche per il riferimento a dimensioni linguistiche e sociolinguistiche distinte. In un

compito escussivo si possono celare, ad esempio, sia aspetti di tipo normativo, che rimandano alla consapevolezza e competenza sociolinguistica, sia molto banalmente questioni legate alla velocità dell'elocuzione, alla coarticolazione dei gesti e così via, fatti cioè riconducibili alle caratteristiche della produzione di parlato (semi)spontaneo. Inoltre, non è chiaro se la differenza di compito abbia lo stesso significato e risponda agli stessi meccanismi in termini fonetici, sociolinguistici e cognitivi per i parlanti nativi e per i parlanti bilingui o di una L2, sebbene con un grado di competenza estremamente elevato.

Prima di illustrare i risultati dello studio, è opportuno indicare sinteticamente i tratti essenziali dell'analisi delle corrispondenze multipla, la tecnica statistica qui impiegata per studiare le relazioni tra i vari descrittori della variazione allofonica.

L'analisi delle corrispondenze multipla[14] (ACM) qui adottata potrebbe essere definita come un'analisi fattoriale per variabili categoriali, poiché appartiene alla famiglia delle tecniche di analisi multivariata che hanno lo scopo di trattare simultaneamente molte variabili, trovando relazioni e riducendo il numero di dimensioni rilevanti. Si tratta delle metodiche di analisi esplorativa dei dati che normalmente non distinguono (ma alcune, come l'ACM, possono anche farlo) tra variabili dipendenti e una o più variabili indipendenti, ovvero non pongono un nesso di causalità forte. Questi metodi hanno piuttosto l'obiettivo di fornire rappresentazioni grafiche sintetiche a partire da matrici di osservazioni di un fenomeno, di norma, descritto da numerose proprietà (per una prima applicazione sperimentale della tecnica in sociolinguistica v. Vietti, 2007, mentre per una trattazione tecnica cfr. Fabbris, 1997; Bolasco, 1999; e ora anche il più accessibile Di Franco, 2006).

In modo molto riassuntivo si può considerare l'ACM come un modo di stabilire la struttura delle relazioni all'interno di una tabella di contingenza attraverso una «rappresentazione grafica delle modalità osservate su uno spazio di dimensionalità minima» (Fabbris, 1997: 251). Poiché si tratta di un metodo più geometrico-grafico che non statistico in senso stretto, si può pensare all'output di questa tecnica

[14] Nelle trattazioni manualistiche in italiano l'aggettivo *multiplo* appare legato ora a una ora all'altra delle due teste nominali: analisi delle corrispondenze *multipla*, oppure analisi delle corrispondenze *multiple*.

come a una mappa strutturale della dipendenza tra i valori delle variabili osservate.
Arrivando dunque alla nostra matrice di dati, questa si compone di 364 righe, le osservazioni del fenomeno, e di quattro colonne, le variabili nominali o categoriche individuate per descrivere il fenomeno. Obiettivo dell'analisi, eseguita con il programma SPSS 17 (funzione *Analyze* > *Dimension Reduction* > *Optimal Scaling*), è dunque di rintracciare possibili co-occorrenze e associazioni tra i valori delle variabili attraverso le 364 osservazioni in modo da offrire una descrizione della variazione allofonica di /R/ nell'italiano di tedescofoni sudtirolesi.

4.3 Analisi delle corrispondenze multiple

Una prima analisi esplorativa compiuta su tre variabili (Varianti, Contesto fonetico e Sillaba) e presentata nel Grafico 1 ha lo scopo di illustrare proprio le caratteristiche distribuzionali dei diversi foni. È necessario però ricordare che, dato l'elevato numero di modalità della variabile Contesti fonetici[15] e il ridotto campione di osservazioni, l'attendibilità e la solidità dell'analisi non sono molto elevate e danno luogo a valori anomali nell'estrazione dei fattori. Pur con queste cautele, la lettura del grafico appare notevolmente informativa.

[15] Sono in ogni caso stati eliminati per questa analisi i contesti con frequenze inferiori a 4 per un *data set* analizzato di 297 occorrenze.

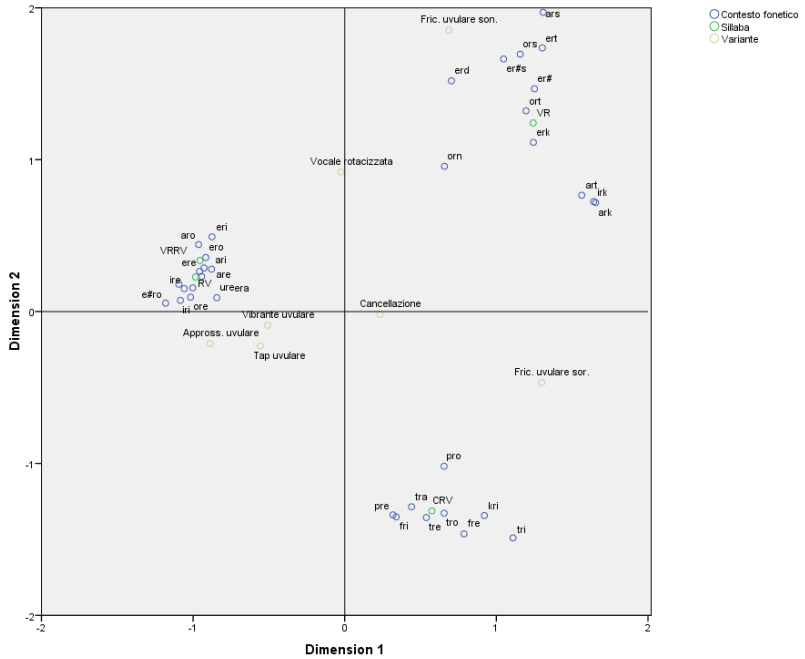


Grafico 1: ACM, distribuzione dei foni (Varianti, Sillaba, Contesto fonetico)

A un primo sguardo appare immediatamente una suddivisione in tre nuvole di punti-modalità delle variabili lungo i due assi individuati. Il ruolo più significativo nella spiegazione della varianza della matrice è svolto dalla variabile Contesti fonetici che risulta in completa sovrapposizione con quella dei tipi sillabici attestati (Sillaba). La dimensione 2 distingue a sinistra i contesti fonetici intervocalici (con al centro RV e VRRV) e a destra quelli caratterizzati da nessi consonantici (CRV e VR). La dimensione 1 scompone ulteriormente i gruppi consonantici in pre- (in alto) e post-consonantici (in basso). Lungo la dimensione 2 (intervocalico – nesso consonantico) i foni si possono ordinare in questo modo: approssimante/tap/vibrante[16], vocale rotacizzata/cancellazione (al centro dell'asse e quindi non fortemente influenzati da questa dimensione), fricativa sonora/fricativa sorda. La dimensione 1 distingue principalmente tra le due fricative: da un lato, quella sonora si associa a contesti RC (non necessariamente sonora), mentre, dall'altro, la sorda mostra una maggiore propensione a occorrere in nessi CR dove la consonante è

[16] Si noti come questi due "oggetti" siano altamente simili sul piano distributivo.

per lo più un'occlusiva sorda (*pre*, *tra*, *kri* ecc.). Da notare ancora (a) che la cancellazione non sembra attirata nell'area di nessuna delle nuvole e si pone anzi vicino all'origine, mostrando così una sostanziale indipendenza da criteri distribuzionali e (b) che la vocale rotacizzata sembra invece esibire associazione ai nessi RC (cfr. in particolare i valori dei contesti *orn*, *art*, *irk*, *ark* sull'asse 2).
Un'interpretazione preliminare di questo piano cartesiano rivela la forte dipendenza degli allofoni di R dalla distribuzione attraverso i contesti e, in particolare, da criteri non di tipo strutturale (posizione di testa o coda sillabica), ma come è prevedibile di tipo lineare come la tripartizione in posizione V_V, _C o C_. In prima approssimazione si potrebbe sostenere che a requisiti energetici come quelli richiesti dalla posizione intervocalica si associ un determinato fascio di allofoni con caratteristiche di alta sonorità ed energia, mentre in presenza di nessi consonantici sembra prevalere l'effetto coarticolatorio della consonante che (a) nel caso di RC può assumere la forma di vocale rotacizzata > vocale rotacizzata con rumore > fricativa sonora, mentre (b) nel caso di CR (spesso con C occlusiva sorda) assume la forma di una fricativa sorda.

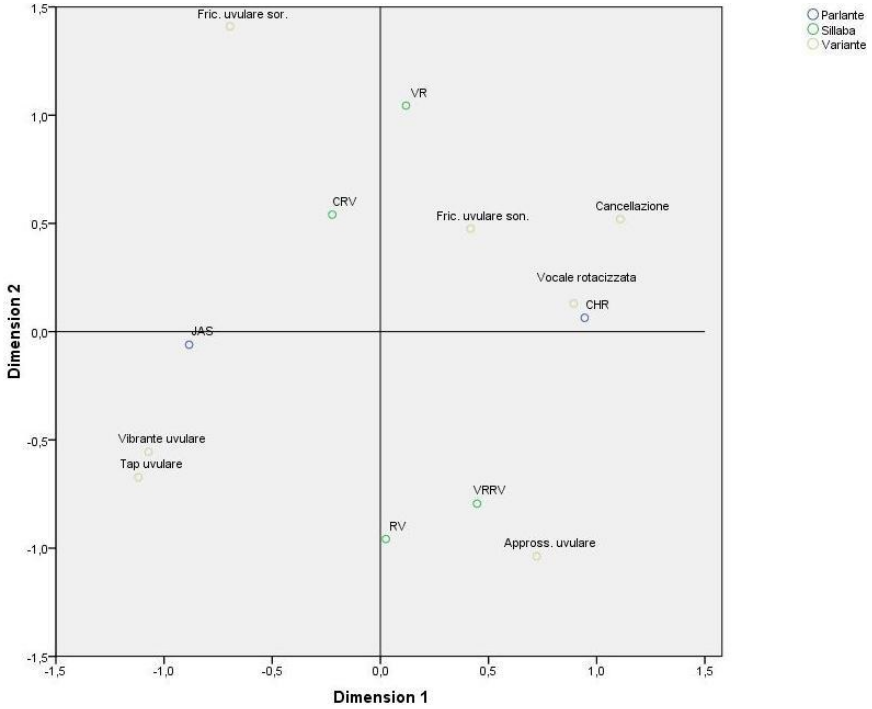


Grafico 2: ACM, Parlante, Sillaba e Variante

Una volta chiarito lo schema generale di distribuzione dei foni è possibile introdurre alcune variabili funzionali che agiscono dall'esterno sul micro-sistema linguistico in esame. Nel Grafico 2 sono rappresentate, sempre in uno spazio bidimensionale, le relazioni tra Variante, Sillaba e Parlante. La variabile Contesto fonetico è stata qui sostituita da Sillaba senza significative perdite di informazione, poiché, come si è potuto constatare nel Grafico 1, la sua posizione nel piano coincideva largamente con quella dei singoli contesti.
Nell'interpretare i fattori individuati nel Grafico 2 è opportuno essere consapevoli che si stanno osservando variabili che agiscono su livelli diversi: mentre il Grafico 1 ci informa sulle relazioni immediate tra le caratteristiche della sequenza prodotta e la selezione dell'allofono, nel Grafico 2 la relazione che possiamo osservare tra il parlante e la distribuzione dei foni è evidentemente esterna e tipicamente sociolinguistica.
La dimensione 2 suddivide i foni in modo quasi perfettamente complementare sulla base del parlante, individuando così due sottosistemi indipendenti l'uno dall'altro, quello di JAS a sinistra e quello di CHR a destra. Sulla base di questo fattore emergono così due varietà distinte di italiano di tedescofoni, contraddistinte nel nostro esempio dalla presenza di due insiemi complementari di foni: tap, vibrante e fricativa sorda per JAS; fricativa sonora, approssimante, vocale rotacizzata e cancellazione per CHR. I suoni con valori più estremi sull'asse 2 – tap (-1,119) e vibrante (-1,072) per JAS, vocale rotacizzata (0,893) e cancellazione (1,109) per CHR – sono evidentemente quelli più fortemente (quasi esclusivamente) associati con un determinato parlante.
Tale partizione è di notevole interesse sociolinguistico poiché costituisce una prima, anche se esile, prova empirica dell'esistenza di almeno due varietà di italiano di tedescofoni altoatesini. Le ragioni che sarebbero alla base di questa differenziazione così marcata riguardo alla distribuzione delle /R/ non è chiara e purtroppo nemmeno investigabile in questo studio. Tuttavia, le ipotesi che si possono formulare si indirizzano soprattutto verso la possibile influenza esercitata dalle varietà di dialetto tirolese (meranese e pusterese) che, nel contatto con l'italiano, trasferirebbero i relativi sotto-sistemi di suoni.

La dimensione 1 invece distingue i suoni all'interno delle due varietà sulla base del criterio distribuzionale. In questo modo si avranno: per la varietà di JAS, tap e vibrante associate al contesto intervocalico e, molto distante e perciò distinta, la fricativa sorda legata alla posizione post-consonantica CRV; mentre nella varietà di CHR i suoni sono ordinati a partire dall'approssimante, alla vocale rotacizzata, alla fricativa sonora e quindi alla cancellazione. La presenza della variabile Parlante rende meno chiaro il pattern di distribuzione emerso nel Grafico 1 schiacciando in particolare VR e CRV, i due contesti con presenza di nessi consonantici, lungo il medesimo asse, ma mettendo altresì in evidenza come per JAS la distribuzione tra tap e vibrante da un lato e fricativa sorda dall'altro sia nettamente legata ai contesti, mentre la maggiore centralità (vicinanza a zero) sulla dimensione 1 della vocale rotacizzata, della fricativa sonora e della cancellazione indicherebbe una distinzione tra questi suoni meno motivata dalla distribuzione contestuale.
Infine, l'aggiunta della variabile Stile all'analisi rappresentata nel Grafico 2 non sembra mostrare schemi significativi e degni di interesse[17], questo forse a causa del fatto che una possibile relazione tra il compito comunicativo, lungo l'opposizione parlato letto vs parlato spontaneo, risulta poco coerente con la struttura relazionale già espressa. Dunque per stabilire se lo stile comunicativo eserciti una qualche azione sugli allofoni delle due varietà di italiano si possono percorre due strade alternative: una è quella di aumentare la dimensionalità dello spazio, l'altra è di eseguire due analisi parallele sulle due varietà di italiano. La prima pista mostra che una terza dimensione non soltanto non porta un contributo significativo alla spiegazione della varianza dei dati, ma aumenta di molto le difficoltà di lettura della rappresentazione grafica.
Le due analisi (delle quali per ragioni di spazio non si possono mostrare i piani) compiute sulle variabili Variante, Sillaba e Stile dei due sotto-campioni di JAS e CHR mettono invece in rilievo un pattern simile legato alla dimensione stilistica che si applica però a oggetti-suoni diversi. Procedendo con ordine, se la prima dimensione in entrambe le analisi si riferisce alla già delineata opposizione distribuzionale tra contesti intervocalici e "consonantici", la

[17] Il contributo di Stile alla varianza della dimensione 1 è 0,084 mentre a quella della dimensione 2 è 0,217.

dimensione ortogonale che potremmo definire "stilistica" oppone lo stile di lettura di parole a quello di frase e del map task. Questi ultimi sarebbero dunque accomunati da valori abbastanza simili sull'asse 2 e piuttosto vicini all'origine, mentre la lettura di parole sarebbe distinta e occuperebbe una posizione distante nel semipiano inferiore[18].
La distinzione tra i due parlanti avviene per quanto riguarda il fono associato a uno stile di lettura di parole in isolamento e, quindi, si presume avvertito come più formale, o atteso secondo la norma. Nel caso di CHR si tratta della fricativa uvulare sorda (si noti *en passant*, fono prototipico nel tedesco standard), mentre in quello di JAS della vibrante uvulare.
Dunque la variabile Stile svolge un ruolo all'interno della rete di relazioni distinguendo da un lato la lettura di parole dagli altri stili elocutivi, e dall'altro il fono ritenuto più adatto a uno stile di parlato letto. Come ultima osservazione va ricordato che la variabile Stile, come tutte quelle qui adoperate, può nascondere al proprio interno fattori e indicatori distinti. Nel passaggio da un compito comunicativo come il Map Task alla lettura di parole in isolamento può cambiare non soltanto l'assetto articolatorio e la velocità di esecuzione, ma anche l'impegno cognitivo che può interessare in modo differente i vari processi cognitivi (attenzione, motivazione ecc.). Infine, queste potenziali fonti di variabilità si innestano evidentemente sulla dimensione socio-comunicativa composta da norme, rappresentazioni e aspettative relative ai comportamenti situazionalmente adeguati. In definitiva, sappiamo molto poco di ciò che si cela sotto le variabili indagate, tuttavia le mappe tracciate, anche se da una notevole "distanza", forniscono un quadro preliminare sulla macro-struttura della variazione allofonica nei suoni /R/ nelle varietà di italiano di tedescofoni in Alto Adige.

5 Conclusioni

Lo studio effettuato sul campione considerato ha parzialmente confermato le precedenti descrizioni dell'italiano di tedescofoni

[18] La posizione centrale di lettura frasi e map task è legata alla maggiore massa (si potrebbe dire frequenza relativa nel campione) rispetto alle occorrenze di lettura di parole.

altoatesini e ha consentito di tratteggiare un più ampio quadro di variazione allofonica anche grazie all'analisi spettrografica.
Inoltre dall'analisi delle corrispondenze multiple sono emersi schemi che evidenziano chiaramente i rapporti tra le variabili considerate. Particolarmente significativi sono quelli che vincolano il fenomeno sul piano sistemico sulla base di criteri distribuzionali, ovvero per contesti fonetici, e quelli che legano la variazione di /R/ al tipo di parlante o allo stile comunicativo sul piano extralinguistico.

Riferimenti bibliografici

Bolasco, S. (1999) *Analisi multidimensionale dei dati. Metodi, strategie e criteri d'interpretazione*. Roma: Carocci.

Canepari, L. (1999) *MaPI. Manuale di Pronuncia Italiana*. Bologna: Zanichelli.

– (2006) *Avviamento alla fonetica*. Torino: Einaudi.

Di Franco, G. (2006) *Corrispondenze multiple e altre tecniche multivariate per variabili categoriali*. Milano: Franco Angeli.

Demolin, D. (2001) 'Some phonetic and phonological observations concerning /R/ in Belgian Frenc'. *Etudes et Travaux* 4, pp. 63-73.

Docherty, G. and P. Foulkes (2001) 'Variability in /r/ production. Instrumental perspectives'. *Etudes et Travaux* 4, pp. 173-184.

Engstrand, O., J. Frid and B. Lindblom (2007) A Perceptual Bridge Between Coronal and Dorsal /r/, in M. Sole, P. Speeter Beddor and M. Ohala (eds) *Experimental approaches to phonology*, Oxford: Oxford University Press, pp. 175-91.

Fabbris, L. (1997) *Statistica multivariata. Analisi esplorativa dei dati*. Milano: McGraw-Hill.

Felloni, M.C. (2006) *Un'indagine sociofonetica a Parma: la realizzazione del fonema /r/ nell'italiano regionale*. Tesi di Laurea Specialistica, Università di Pavia.

Foulkes, P. and G. Docherty (2001) 'Variation and change in British English /r/'. *Etudes et Travaux* 4, pp. 27-43.

Klein, K. and L. Schmitt (1969) *Tirolischer Sprachatlas*. Innsbruck: Tyrolia-Verlag.

Ladefoged, P. and I. Maddieson (1996) *The sounds of the world's languages*. Oxford: Blackwell.

Maturi, P. (2006) *I suoni delle lingue, i suoni dell'italiano*. Bologna: il Mulino.

Mioni, A. (1990) La standardizzazione fonetico-fonologica a Padova e Bolzano (stile di lettura) in M. Cortelazzo and A. Mioni (eds) *L'italiano regionale*, Roma: Bulzoni, pp. 193-208.
– (2001) L'italiano nelle tre comunità linguistiche tirolesi, in K. Egger and F. Lanthaler (eds) *Die Deutsche Sprache in Südtirol*, Wien: Folio, pp. 65-76.
Nespor, M. (1993) *Fonologia*. Bologna: il Mulino.
Recasens, D. (1991) 'On the production characteristics of apicoalveolar taps and trills'. *Journal of Phonetics* 19, pp. 267-80.
Romano, A. (in press) A contribution to the study of phonetic variation of /r/ in French and Italian linguistic domains, in H. Van de Velde, R. van Hout and D. Demolin (eds) *'r-atics. Sociolinguistic, phonetic and phonological characteristics of /r/*.
Schiller, Niels O. (1998) The Phonetic Variation of German /r/, in M. Butt and N. Fuhrhop (eds) *Variation und Stabilität in der Wortstruktur. Untersuchungen zu Entwicklung, Erwerb und Varietäten des Deutschen und anderer Sprachen*, Hildesheim: Olms, pp. 261-87.
Schmidt, S. (1999) *Fonetica e fonologia dell'italiano*. Torino: Paravia.
Simpson, A. (1998), Phonetische Datenbanken des Deutschen in der empirischen Sprachforschung und der phonologischen Theoriebildung'. *Institut für Phonetik und digitale Sprachverarbeitung, Universität Kiel, Arbeitsberichte* 4, pp. 1-233.
Stuart-Smith, J. (2007) A sociophonetic investigation of postvocalic /r/ in Glaswegian adolescents, in J. Trouvain, W. J. Barry (eds) *Proceedings of the XVIth International Congress of Phonetic Science,* Saarbrucken: Dudweiler, Pirrot GmbH, pp. 1449-52.
Tonelli, L. (2002) *Regionale Umgangssprachen*. Padova: Unipress.
Van de Velde, H. and R. Van Hout (1999) The pronunciation if (r) in Standard Dutch, in R. van Bezooijen and R. Kager (eds) *Linguistics in the Netherlands 1999*, Amsterdam: Benjamins, pp. 177-88.
Vietti, A. (2007) *Contesti d'uso in repertori linguistici complessi. Tentativi di descrizione multidimensionale dei dati della* Survey Ladins, in G. Iannàccaro and V. Dell'Aquila (eds) *Survey Ladins*, Atti del convegno (Vigo di Fassa 14-16.9.2006) «Mondo Ladino», 31, pp. 239-66.
Vietti, A. and L. Spreafico (2008) 'Phonetic variation of /r/ in a language contact situation', Poster presentato a *LabPhon11*, Wellington, New Zealand 30 June-2 July 2008.

Vietti, A., L. Spreafico and A. Romano (in press) Tempi e modi di conservazione delle /r/ italiane nei *frigoriferi* CLIPS, in *Atti del 5° convegno nazionale AISV, Zurigo 4-6 febbraio 2009*.
Wiese, R. (2001) 'The unity and variation of German /r/'. *Etudes et Travaux* 4, pp. 11-26.

The pragmatic ground of *flores rhetoricae*[*]

Carlotta Viti
Universität Jena, Universität Zürich
carlotta.viti@klphs.uzh.ch

Abstract

The present contribution provides comparisons between recent studies of pragmatics and the ancient tradition of rhetoric. In particular, we observe the naturalness of rhetorical figures and poetic style in spoken discourse. This claim is supported by two main pieces of evidence. First, the written texts of Latin and Ancient Greek were profoundly different from the written texts of their daughter languages, since in the former orality had a much more prominent role. Second, certain rhetorical figures – particularly metaphor – have been proven to be by no means limited to writing, but rather as being relevant phenomena of everyday spoken conversation. The same may hold true for other rhetorical figures of meaning as well as of form.

1 Introduction

The early Indo-European languages are often neglected in studies of spoken communication, owing to the fact that they are available only in written documents and often display patterns that are at odds with syntactic principles of constituency, adjacency, government, etc. as well as with pragmatic principles of clarity that are normally observed in natural languages. These seemingly aberrant patterns are usually considered as being artificial, literary creations, and are often relegated in the domain of rhetorical figures (*flores rhetoricae*), cf. Curtius (1963); Lausberg (1990); Üding and Steinbrink (1994); Maurach (1995), etc. (§2).
However, the earliest documents of languages such as Ancient Greek, Latin, Vedic, Hittite, etc. clearly reflect a 'pragmatic communicative mode' in Givón's (1979) terms, according to which clauses show a

[*] This paper is part of the project 'Die Informationsstruktur in älteren indogermanischen Sprachen' funded by the Deutsche Forschungsgemeinschaft and performed at the Friedrich-Schiller University of Jena under the direction of Prof. Dr. Rosemarie Lühr. I am grateful to Frau Lühr and to my colleague Dr. Susanne Schnaus for their help and support.

loose, appositive, and mainly paratactic linkage, where coreference is often not established by syntactic means, but rather by comprehension of the whole context, as well as of the background knowledge of the speech act participants (§3). Moreover, some rhetorical figures such as metaphor have been convincingly shown not to be limited to literary texts, but rather to be pervasive of daily spontaneous conversations (Lakoff and Johnson, 1980) (§4.1). It is conceivable that the same pragmatic grounding may be also identified in other *flores rhetoricae*, such as accumulation, alliteration, anacoluthon, asyndeton, hyperbaton, etc. (§§4.2, 5).

2 Discourse and rhetorical figures in antiquity

Traditionally, rhetorical figures are considered as being stylistic devices that may be used to embellish a discourse. They represent a manifestation of *elocutio*, in Ancient Greek *léxis*, the 'art of well-speaking, style' (Ger. *sprachlicher Ausdruck*), which is one of the five basic aspects of an oration[1].

Well-speaking may be identified through 'stylistic virtues' (*virtutes elocutionis*), the main being correctness, clarity, and ornateness. Correctness (Lat. *latinitas*, Gr. *hellenismós*, Ger. *Sprachrichtigkeit*) implies that one should speak in well-formed structures, chosen according to the common use, and avoid any idiosyncratic or aberrant expressions, such as barbarisms, solecisms, etc. Clarity (Lat. *perspicuitas*, Gr. *saphaneîa*, Ger. *Deutlichkeit*) indicates a direct, straightforward expression, which avoids any ambiguity or obscurity. Finally, ornateness (Lat. *ornatus*, Gr. *kataskeuḗ,* Ger. *Redeschmuck*) represents the set of devices which make a speech more adorned[2]. Rhetorical figures properly pertain to the domain of ornateness.

[1] The other parts are *inventio* 'invention', that is, finding something to say (Gr. *heúrēsis*, Ger. *Erfinden*), *dispositio* 'arrangement' of the various argumentations in a coherent sequence (Gr. *táxis*, Ger. *Anordnen*), *actio* 'performance', i.e. the set of gestures that may increase the effect of the discourse (Gr. *hupókrisis*, Ger. *körperliche Beredsamkeit*) and *memoria* 'memory', the use of mnemonic techniques to remember the various points of a speech (Gr. *mnḗmē*, Ger. *Gedächtnis*).

[2] The list of *virtutes elocutionis* may be enlarged, including, for example, *decorum* and *evidentia*. *Decorum* 'propriety' (Gr. *prépon*, Ger. *Angemessenheit*) implies that words must be chosen according to the subject matter. *Evidentia* 'evidence' (Gr.

Not all *virtutes elocutionis* have the same importance in the literary and grammatical tradition of antiquity. On the one hand, there is a common consensus that *latinitas* and *perspicuitas* must be observed in all types of genres and texts. On the other, not all authors agree that *ornatus* is to be pursued or even studied.

The universal appreciation of *latinitas* and *perspicuitas* may be related to the fact that these properties closely correspond to the fundamental principles underlying human conversation, which have been acknowledged since early studies of pragmatics, such as Grice's 'conversational maxims' (1989). In particular, the Latin term *perspicuitas* is faithfully mirrored by the Gricean maxim of Manner, which runs as follows: 'be perspicuous', and which has the following corollaries: 'avoid obscurity of expression; avoid ambiguity; be brief (avoid unnecessary prolixity); be orderly'. More generally, the other Gricean maxims of conversation, that is, Quality ('make your contribution one that is true'), Quantity ('make your contribution as informative as is required'), Relation ('be relevant') may also be implied by the Latin *perspicuitas*, as can be seen in the following definition of this virtue of speech given by Quintilian (first century AD):

> Nobis prima sit virtus perspicuitas, propria verba, rectus ordo, non in longum dilata conclusio, nihil neque desit neque superfluat: ita sermo et doctis probabilis et planus imperitis erit. […] Nam si neque pauciora quam oportet neque plura neque inordinata aut indistincta dixerimus, erunt dilucida et neglegenter quoque audientibus aperta.
> [For my own part, I regard clearness as the first essential of a good style: there must be propriety in our words, their order must be straightforward, the conclusion of the period must not be long postponed, there must be nothing lacking and nothing superfluous. Thus our language will be approved by the learned and clear to the uneducated. […] For if what we say is neither less nor more than is required, and is clear and systematically arranged, the whole matter will be plain and obvious even to a not too attentive audience] (Quint. *Inst.Or.* 8.2.22)

---

enárgeia, Ger. *Augenscheinlichkeit*) implies that the matter is described as if occurring before one's eyes.

Quite differently, the judgment on *ornatus* is controversial. Sophists such as Gorgias from Leontini (V century BC), to whom the introduction of rhetorical figures is traditionally ascribed, were notorious for their unscrupulous speech, capable of making people believe one thing and also the opposite of it. This practice was especially condemned by Roman society, where oratory was always related to civil duties and politics. Cato the Elder (234-149 BC), for example, the first Roman author of whom at least some oratory excerpts are preserved, contested the concerns of Greek rhetoric. His famous ditto *rem tene, verba sequentur* 'hold the matter, the words will follow' implies a major importance of the matter (*res*) with respect to the style (*verba*).

Cato's thoughts are continued by Mark Anthony (143-87 BC), the grandfather of the homonymous triumvir, author of a (lost) handbook of oratory entitled *De ratione dicendi*. Mark Anthony – together with his contemporary Lucius Licinius Crassus – is one of the protagonists of Cicero's *De Oratore*. Crassus was the advocate of a cultivated oratory, performed by accomplished connoisseurs of Greek literature, where the *elocutio* played a major role. Conversely, Mark Anthony's oratory aimed to be free of the Greek models, considered the raw facts and gave only little attention to stylistic matters. The Greek influence increases with time on Latin authors, so that the oratory of Licinius prevailed over that of Mark Anthony; Cicero himself supported a cultivated style, where rhetorical figures were attentively calibrated. However, in all classical treatises on style (Plato's *Gorgias*, V BC; Aristotle's *Rhetoric*, IV BC; Cicero's *De Oratore*, *Orator*, *Brutus*, the anonymous *Rhetorica ad Herennium*, I BC; Longinus' *Perì Hýpsous*, I AD; Quintilian's *Institutio Oratoria*, I AD; Hermogene's *Perì idéōn*, I AD, etc.), rhetorical figures were considered as being deviations from the correct and direct way of speaking. Especially in the Latin tradition, an excessive use of rhetorical figures was judged negatively as evidence of bad taste, and even of insincere or non-honest speech.

Rhetorical figures are found more often in poetry than in prose, so that a sort of continuum of deviation and artificiality has been recognized between a regular poetic feature such as meter and a frequent poetic feature such as rhetorical figures. It must be emphasized that the classical grammatical tradition had a normative rather than a descriptive character – which represents the main difference between classical rhetoric and modern pragmatics. Accordingly, scholars

continued for centuries to look at rhetorical figures as poetic or literary entities rather than as a manifestation of actual language in use.

3 The oral character of ancient texts

Although Latin and Ancient Greek are only available through written documents (which often show gaps, interpolations, and corruptions in their philological tradition), they present clear features of oral communication. Most of the classical texts, especially in the earliest periods, were composed to be pronounced aloud in front of a public rather than to be read in private. This applies not only to orations or theatrical pieces, of which oral performance is evident, but also to the historical or narrative genre. Herodotus' *Historiae*, for example, were openly recited in Athens. In the philosophical genre, Socrates did not write anything, and Plato's dialogues allude to an interactive way of teaching, which analyzes complex concepts by means of gradual questions and answers between the teacher and his disciple. Even poetry, which in principle seems to imply an individual performance, was often brought to social fruition in antiquity; the Homeric poems, with their long oral tradition, are the clearest example of this.

The oral character of these ancient texts also has more structural manifestations, in various areas of grammar. First, word order is syntactically free and pragmatically determined, that is, various word order arrangements may be exploited to convey different types of information. Although the specific representation of topic, focus, etc. in the various constructions of a text is not without controversy (for Latin, cf. Panhuis, 1982; Devine and Stephens, 2006; for Ancient Greek, cf. Dover, 1960; Dik, 1995; Matic, 2003; Lühr, forthcoming), all scholars agree that the word order flexibility of the ancient languages is conditioned by discoursive principles, and that much of its interpretative difficulty is due to the loss of prosodic information. In these languages, a word or group of words is usually moved towards the beginning of the sentence if it is prominent or important, that is, if it recurs several time in the text. Meillet, for example, compares the fixed word order in French *Pierre bat Paul* to the multiple solutions of Latin *Petrus Paulum caedit*, *Paulum Petrus caedit*, *caedit Petrus Paulum*, *caedit Paulum Petrus* etc.:

> L'ordre n'est pas indifférent; il sert à indiquer certaines nuances: suivant qu'on met en avant *Petrus* ou *Paulum*, on attire l'attention sur l'un ou l'autre mot; mais l'ordre n'indique en rien le rôle grammatical des mots. (Meillet, 1921: 147)

Second, clause linkage is often performed by means of asyndeton or of implicit devices such as particles. Ancient Greek especially displays an extensive – and often elusive – use of sentence particles (cf. Denniston, 1950; Bonifazi, 2009), but even in Latin sentence particles are more frequent than the conventional conjunctive, disjunctive and adversative connectors (cf. Kroon, 1995; 2009; Krylová, 2009). In the classical languages, and more generally in the ancient Indo-European languages (cf. Lühr, 2009), particles were often employed to express degrees of evidence or commitment, to organize the text into different sections, or to maintain coherence among the various argumentations. Such exigencies are more cogent in oral speech than in writing, where the reader can easily recover the previous information. Both *nam* and *enim*, for example, were traditionally considered as being causal adverbs, to be translated in English 'for', German *denn*, French *car*, Italian *infatti*. Caroline Kroon, however, convincingly shows that a pragmatic difference exists between *nam* and *enim*: the former generally indicates a subsidiary situation, with the possible side-effect of causality, while the latter has an additional meaning of appeal to interpersonal consensus and operates at the interactional level of discourse, similarly to English *of course*, *you know*, *indeed*, *clearly*, *you will agree*, *mark you*. Similar expressions are categorized as 'discourse markers' (Schiffrin, 1987), and appear in the oral informal register of English.

Third, it is typical of the early text of Latin and Ancient Greek, as well as of the other old Indo-European languages, that certain constituents are not syntactically connected to the rest of the sentence. When the isolated constituent is inflected in the nominative, we have the *nominativus pendens*, which is often interpreted as an *anacoluthon* (cf. Havers, 1922). When, much more frequently, it is found in an oblique case – ablative in Latin, genitive in Ancient Greek – we have an absolutive construction. Absolutive constructions are synchronically interpreted as expressing the temporal circumstances of the main clause situation, e.g. *Cethegus, recitatis litteris, repente conticuit*

'When the letter had been read out, Cethegus suddenly fell silent' (Cic. *Cat.* 3.10). Diachronically, however, the origin of these detached phrases is to be related to the oral discourse (cf. Ramat, 2005), which privileges a loose connection among juxtaposed elements over condensed, syntacticized syndetic devices. As anticipated, all these constructions are manifestations of the 'pragmatic communicative mode', which Givón (1979) opposes to the 'syntactic communicative mode'. Owing to this, pragmatic studies of various aspects of the grammar of the ancient languages have appeared in the last years. The Jena-Project on the Information Structure of the old Indo-European language investigates, in particular, the change from more pragmatic-oriented languages to more syntactic-oriented languages.

The pragmatic, discoursive motivation of structural phenomena is more evident if one considers syntactic constructions, as in the examples reported above, as syntax is more transparent and prone to variation and change than morphology. Underlying pragmatic functions, however, also appear in the domain of morphology. The classical languages have a rich morphology, in both noun and verb system. Since morphological information is obligatory, it is also redundant, in the sense that, for example, finite verbs must always be inflected for tense, mood, person and number, and nouns must always be inflected for gender, number and case, even though their temporal, modal and referential properties may be retrieved from the context. Owing to this redundancy, morphological information may be exploited for non-grammatical purposes. For example, tenses often have non-literal uses (*praesens pro praeterito*, *praesens pro futuro*, etc.), which express the speaker's subjective perspective on the events. In Homeric Greek even the morphological marking of a marginal domain such as the dual may have a discoursive motivation (cf. Viti, 2008). It would be strange if the ancient texts, which are so pervasively characterized by features of spoken communication, were to present something as artificial as rhetorical figures have traditionally been thought to be.

4 Naturalness of rhetorical figures

4.1 Figures of meaning

The number and type of rhetorical figures vary according to the different traditions of literary criticism. Quintilian (first century AD)

discusses matters of style in the VIII book of his *Institutio Oratoria* (*Hinc enim iam elocutionis rationem tractabimus, partem operis, ut inter omnis oratores convenit, difficillimam* 'For I have now to discuss the theory of style, a subject which, as all orators agree, presents the greatest difficulty'), and presents a list of rhetorical figures in section VIII.6, such as metaphor, simile, synecdoche, metonymy, antonomasia, onomatopoeia, catachresis, metalepsis, epithet, allegory, irony, hyperbaton, and hyperbole. The *Rhetorica ad Herennium* lists a much higher number of figures. In general, rhetorical figures may concern either expressions of meaning, e.g. the use of certain lexemes outside their proper literal domain, or expressions of form, where attention is given to the sound of a structure and to its peculiar morpho-syntactic properties[3]. Not all of them, however, have been considered as being equally artificial.

Metaphor was the first rhetorical figure acknowledged as having a certain naturalness. This claim is usually related to the essay *Metaphors we live by*, published in 1980 by George Lakoff, a linguist, and Mark Johnson, a philosopher. The two American scholars observe that metaphor is not only a literary expression that may be used to adorn our style. Quite differently, it is frequently used in our everyday conversational transactions, and may even shape our conceptualization of extra-linguistic world and personal experience. We report the incipit of this successful book:

> Metaphor is for most people a device of the poetic imagination and the rhetorical flourish – a matter of extraordinary rather than ordinary language. Moreover, metaphor is typically viewed as characteristic of language alone, a matter of words rather than thought or action. For this reason, most people think they can get along perfectly well without metaphor. We have found, on the contrary, that metaphor is pervasive of everyday life, not just in language but in thought and action. Our ordinary conceptual system, in terms of which we both think and act, is fundamentally metaphorical in nature. (Lakoff and Johnson, 1980: 3)

[3] This distinction does not coincide with the traditional distinction drawn between Figures of Speech (*figurae dictionis*, Gr. *schemata léxeos*) and Figures of Thought (*figurae sensus*, Gr. *schemata diánoias*), since the latter refers more to *inventio* than to *elocutio* (cf. *Ad Her*. 4.12.18; Quint. 9.1.15-21).

For example, the metaphor 'argument is war' underlies many common expressions such as 'your claims are *indefensible*', 'he *attacked every weak point* in my argument', 'his criticisms were right *on target*', 'I *demolished* his arguments', etc. (p. 4). As can be seen, metaphor presents an abstract concept by means of a concrete image, and this determines its 'vivid' representation. Lakoff and Johnson (1980) have the merit of having brought the issue of the naturalness of metaphor to the attention of the great public. Their observations, however, were anticipated by the classical rhetorical tradition. Quintilian, for example, starts his discussion of rhetorical figures from metaphor, which he considers as being 'the most beautiful figure' (*pulcherrimus tropus*), and states that it is also 'the most frequent' (*frequentissimus*) figure in ordinary discourse:

> Incipiamus igitur ab eo qui cum *frequentissimus* est tum longe pulcherrimus, translatione dico, quae μεταφορά Graece vocatur. Quae quidem *cum ita est ab ipsa nobis concessa natura ut indocti quoque ac non sentientes ea frequenter utantur*, tum ita iucunda atque nitida ut in oratione quamlibet clara proprio tamen lumine eluceat.
> [Let us start, however, with that species of trope, which is both *the most common* and by far the most beautiful, I mean, with translatio, which the Greeks call μεταφορά. *Metaphor is not only so natural to us, that the illiterate and others often use it unconsciously*, but is so pleasing and ornamental, that, in any composition, however brilliant, it will always make itself apparent by its own lustre] (Quint. *Inst. Or.* 8.6.4)

Here Quintilian argues that metaphor is given by nature (*ab ipsa nobis concessa natura*), so that even uncultivated people can make use of it. He mentions some common metaphors in rural language (*rusticitas*), for example when farmers use the noun *gemma* 'bud', literally 'jewel', when referring to the vines, or when they say that 'the corn thirsts' (*sitire segetes*) and that 'the crops suffer' (*fructus laborare*). Metaphors are very often used due to the lack of a proper term (*indigentia* or *necessitate*): 'from necessity we say that a man is hard or rough (*durum hominem aut asperum*), because there is no proper term for us to give to these dispositions of the mind'. Other times, however, a metaphorical expression may also be given if a more literal term exists, in order to strengthen a concept: 'we say that a man is

inflamed with anger (*incensum ira*), burning with desire (*inflammatum cupiditate*), and has fallen into error (*lapsum errore*), with a view to significance or force of expression, for none of these phrases would be more significant in its own words than in those adopted metaphorically'. Also in these cases metaphors use concrete images to express abstract situations.

4.2 Figures of form

Both Quintilian and Lakoff and Johnson provide observations on similitude and metonymy, similar to those made on metaphor, as all these figures represent a non-literal meaning of an expression. Metaphor and similitude are especially related, differing only in a comparative marker, which is present in similitude (*this man is like a lion*) and absent in metaphor (*this man is a lion*). In general, figures of meaning are considered as being expressions of everyday language more easily than figures of form. Still, a certain naturalness has also been occasionally acknowledged to figures of form in modern studies of pragmatics. For example, lists or enumerations have been shown to be discoursive strategies, commonly used in shopping, in providing evidence in or against an argument, in comparing costs and benefits, etc. (Gilbert, 1989; Jefferson, 1990; Schiffrin, 1994). These strategies clearly match with the classical figures of summary such as *accumulatio*, *anacephalaoesis*, *complexio*, *enumeratio*, *epanodos*, *epiphonema*, *synathroesmos*, etc. which are described in the *Rhetorica ad Herennium* and in other classical and medieval treatises of rhetoric (cf. Calboli Montefusco, 2003).

It has been not adequately recognized, however, that *all* rhetorical figures, even figures of form, have a discoursive basis. Of course, they may be exploited to enhance the beauty and complexity of a certain *modus dicendi*, but their original motivation is pragmatic in nature. This may be seen even in those figures that seem to disrupt the plain and orderly style recommended by the virtues of *latinitas* and *perspicuitas*, such as *anastrophe*, *hysteron proteron*, *hypallage*, *hysterologia*, *parenthesis*, *epergesis*, *synchusis*, etc. It must be observed that hyperbaton *et similia* are different from the word order freedom described in §3, although they are of course related to it. Word order freedom simply implies various sequences of different constituents, while hyperbaton entails disruption in the arrangement of

words inside *the same* constituent. Italian and Spanish, for example, have word order freedom but do not allow disruption of a phrase. These figures may challenge certain syntactic theories, notably those matured in the structuralist or generativist research traditions, which are mainly based on the modern languages of Europe. From the formal point of view, Standard Average European is configurational, that is, it may be adequately described by means of various syntactic configurations such as NP, VP, AP, PP, etc. By definition the elements constituting a phrase are adjacent to each other, cannot be interrupted by heterogeneous material, and if they move they do so all together, according to the so-called 'pied-piping' in generative parlance. Generative grammar attempted to make these configurations explicit, and ascribed them to the I(nternalized)-language, that is, the language which each speaker unconsciously possesses and which children build according to some pre-wired, universal mechanisms. Even outside the generative framework, principles of constituency have been proven to be relevant, since they are typical or at least much more frequent than alternative patterns.

The rationale of the figures of disorder is that in the ancient Indo-European languages even one member of a phrase could be put into focus. Pherecydes of Athens, for example, says: *tòn mèn* ***dexiòn*** *hupodeîtai* ***póda****, tòn dè aristeròn epilḗthetai* 'he put a shoe on his **right foot** but forgot his left one'. In the Ancient Greek text, the adjective *dexión* 'right' is separated from its head noun *póda* 'foot'[4], while in the English translation they are adjacent, and any intervening element would lead to ungrammaticality. Here the contrastive focus is placed on the adjective *dexión* 'right' rather than on the whole phrase *dexión ... póda* 'right foot', since the head noun 'foot' is also implied in the other constituent *aristerón* 'the left one'. The principle 'put the contrastive focus first' overrules in the classical languages the principle according to which words that are conceptually related, such as the members of a phrase, are to be put in an adjacent position. This is a confirmation of the loose linkage which exists even within the members of the same phrase in the ancient languages. It seems that hyperbaton in Ancient Greek and Latin may express the same

[4] We do not consider the intervention of particles (*tòn mèn dexión … tòn dè aristerón*) as being an instance of hyperbaton, since particles are not content-words and usually have a fixed position.

focusing function which in English and especially in French are associated to cleft- or pseudo-cleft sentences.

5 Naturalness of formulas and of meter

Word order is especially deviated in meter, which is certainly the stylistic device most credited to be artificial and apt to manipulability for aesthetic purposes. This assumption is largely related to the name of Milman Parry. On the basis of the observation of metrical procedures in the oral poetry of former-Yugoslavia, Parry (1930) viewed the Homeric language as traditional and formulaic, and showed how a metrical unit such as the 'formula' determines the language of Homer. A formula is a more or less established pairing between one meaning and more than one word; in Parry's terms, it is «a group of words which is regularly employed, under the same metrical conditions, to express a given essential idea» (p. 272). Parry observes that, when the same referent is denoted by two or more different epithets (e.g. *polútlas dîos Odusseús* 'much-suffering godlike Odysseus' and *polúmētis Odusseús* 'many-minded Odysseus'), the epithets differ in metrical structure, so that the poet chooses either the one or the other on the basis of purely metrical considerations.
Subsequently, however, the examination of other types of oral poetry, for example in India or in Finnish territories, suggested to give a greater importance to the meaning component ('essential idea') than to the form component ('a group of words') in the definition of the formula provided by Parry. Scholars presented a reading of Homeric diction that was more independent from the metrical and philological tradition and more permeable to semantic and pragmatic analyses. We limit ourselves to quoting three studies that have been particularly significant with regard to this. The first is a paper by Paul Kiparsky, presented at a meeting on oral literature, which took place at the University of Michigan in November 1974 and was directly inspired by the theories of Milman Parry and by his scholars Albert M. Lord and Joseph A. Russo (cf. Kiparsky, 1976). The second is a more recent book by Egbert Bakker, published in 1997, entitled *Poetry in speech, orality and Homeric discourse.* The third is a series of studies by an Italian scholar of Greek literature, Vincenzo Di Benedetto, whose work culminated in the volume *Nel laboratorio di Omero* in 1994.

The definition of the formula given by Kiparsky ('a ready-made surface structure') relates the rigidity of oral poetry to the bound expressions of everyday language, e.g. 'livelong day', 'foregone conclusion', etc. The difference between such bound expressions and Homeric formulas is one of degree, not of nature. Kiparsky's pioneering insights may be complemented by more recent studies of Construction Grammar (cf. Goldberg, 1995), a theory inspired by the Functionalist research tradition, and therefore far from Kiparsky's personal orientation, which is closer to the generativist model. Construction Grammar is rooted in the assumption that two or more different forms also have a different function, when their alternation cannot be explained on purely structural grounds. The fact that different theoretical frameworks may be compatible with Kiparsky's claims supports their general validity.

Bakker's book makes an interesting connection between the metrical and formulaic language of Homer and contemporary discourse analysis, whereby a metrical unit such as the *colon* corresponds to one piece of information. This claim is especially influenced by the studies carried out by Chafe (1994), according to whom the discourse flow is regulated by the principle of 'one chunk of new information at a time'. Different pieces of information are often separated by sentence particles, which function as 'boundary markers' that set off 'discourse segments against each other' (Bakker, 1997: 51). Accordingly, meter is not a factor that is super-imposed to the language of Homeric poets, but is rather the result of routinization of particularly frequent structures:

> It is true enough that verse necessitates certain patterns and regularities, but it is equally true […] that meter is not an external constraint, independent of discourse. In fact, meter is conditioned by certain phrases just as much as it is itself a conditioning factor, and we might see its relation to the stylized speech units of epic discourse as reciprocally defining: phrases confer their rhythmical and prosodic properties on meter, after which they become metrical, part of the system that is meter and occurring within the contexts created by that system. (Bakker, 1997: 187)

Di Benedetto (1994) illustrates how the use of a formula at a certain point of the discourse may remind the hearer of a previous episode in which the same formula appeared. Thus, the description of Hector's funeral at the end of the *Iliad* echoes the preceding description of Patroklus' funeral, since the murder of Patroklus' at the hands of Hector is avenged through Hector's death. Similarly, in Virgil's *Aeneid*, the final killing of Turnus at the hands of Aeneas is reminiscent of a previous episode in which Pallas, Aeneas' friend, was killed by Turnus. The use of the same formulas warns the hearer that Hector will end up in the same way as Patroklus, and Turnus in the same way as Pallas. Thus, formulas may be exploited to remind and to anticipate, as signals of memory and of imagination, as well as of meta-literary quotations.

This may be underlined by alliteration. A sequence of sounds may also be used to make the hearer become aware of the importance of an episode. This occurs, for example, in a discourse by Ennius (239-169 BC), the author of the earliest Latin poem in hexameters: *o Tite, tute, Tati, tibi tanta, tyranne, tulisti* 'Titus Tatius, you attracted so great disgraces to yourself'. Alliteration implies that a structure must be given special attention, and is a pervasive feature of early Latin texts. The possibility of a non-structural motivation for meter, the greatest manifestation of literacy and extraordinary expression in the *opinio communis*, may provide independent evidence for identifying a discoursive motivation also in rhetorical figures.

6 Conclusions

In this short paper we have observed that ancient languages may enrich the synchronic approach of pragmatic studies with their long diachronic record, and that many principles of modern pragmatics were foreseen by the classical rhetorical tradition. Even constructions which are viewed as artificial and detached from ordinary language may be discoursively and cognitively explained. They may well be called *flores rhetoricae*, but it must be kept in mind that these flowers have a pragmatic ground.

References

Bakker, E.J. (1997) *Poetry in speech, orality and Homeric discourse*. Ithaca-London: Cornell University Press.

Bonifazi, A. (2009) The pragmatic meaning of some discourse markers in Homer, in E. Rieken and P. Widmer (2009) (eds) *Pragmatische Kategorien. Form, Funktion, Diachronie*, Wiesbaden: Reichert, pp. 29-36.

Calboli Montefusco, L. (2003) (ed.) *Papers of Rhetoric, Atti del Convegno Internazionale 'Dictamen, poetria and Cicero: Coherence and Diversification'* (Bologna, 10-11 Maggio 2002), Roma: Herder.

Chafe, W. (1994) *Discourse, consciousness, and time*. Chicago-London: The University of Chicago Press.

Curtius, E.R. (1963)[4] *Europäische Literatur und lateinisches Mittelalter*. Bern-München: Francke.

Denniston, J.D. (1950)[2] *The Greek particles*. London: Duckworth.

Devine, A.M. and L.D. Stephens (2006) *Latin Word Order. Structured Meaning and Information*. New York: Oxford University Press.

Di Benedetto, V. (1994) *Nel laboratorio di Omero*. Torino: Einaudi.

Dik, H. (1995) *Word order in Ancient Greek. A pragmatic account of word order variation in Herodotus*. Amsterdam: Gieben.

Dover, K.J. (1960) *Greek Word Order*. Cambridge: Cambridge University Press.

Gilbert, B. (1989) *La série énumérative. Étude linguistique et stylistique s'appuyant sur dix romans français publiés entre 1945 et 75*. Geneva-Paris: Libraire Droz.

Givón, T. (1979) *On understanding grammar*. New York: Academic Press.

Goldberg, A. (1995) *Constructions. A Construction Grammar approach to argument structure*. Chicago-London: The University of Chicago Press.

Grice, H.P. (1989) *Studies in the way of words*. Cambridge, MA: Harvard University Press.

Havers, W. (1922) 'Der sog. Nominativus pendens'. *Indogermanische Forschungen* XLIII, pp. 207-57.

Jefferson, G. (1990) List construction as a task and interactional resource, in G. Psathas (ed.) *Interaction competence*, Washington DC: University Press of America, pp. 63-92.

Kiparsky, P. (1976) Oral Poetry: Some Linguistic and Typological Considerations, in B. Stolz and R. Shannon (eds) *Oral Literature and the Formula*, Ann Arbor: Center for Coordination of Ancient and Modern Studies, pp. 73-106.

Kroon, C. (1995) *Discourse particles in Latin*. *A study of* nam, enim, autem, vero *and* at. Amsterdam: Gieben.

– (2009) Latin linguistics between Grammar and Discourse. Units of analysis, levels of analysis, in E. Rieken and P. Widmer (2009) (eds) *Pragmatische Kategorien. Form, Funktion, Diachronie*, Wiesbaden: Reichert, pp. 143-58.

Krylová, B. (2009) Zur Diachronie der lateinischen Diskuspartikeln: methodologische Überlegungen, in E. Rieken and P. Widmer (eds) *Pragmatische Kategorien. Form, Funktion, Diachronie*, Wiesbaden: Reichert, pp. 159-72.

Lakoff, G. and M. Johnson (1980) *Metaphors we live by*. Chicago: University of Chicago Press.

Lausberg, H. (1990)[3] *Handbuch der literarischen Rhetorik*. Stuttgart: Steiner.

Lühr, R. (2009) P2 Partikeln in indogermanischen Sprachen, in E. Rieken and P. Widmer (eds) *Pragmatische Kategorien. Form, Funktion, Diachronie*, Wiesbaden: Reichert, pp. 173-86.

Lühr, R. (forthcoming) Old Greek Information Structure, in A. Steube (ed.) *The Discourse Potential of Underspecified Structures*, Berlin-New York (Language, Context, and Cognition).

Matić, D. (2003) 'Topic, focus, and discourse structure. Ancient Greek word order.' *Studies in Language* XXVII, pp. 573-633.

Maurach, G. (1995) *Lateinische Dichtersprache*. Darmstadt: WBG.

Meillet, A. (1921) *Linguistique historique et linguistique générale*. Paris: Champion.

Panhuis, D. (1982) *The Communicative Perspective in the Sentence: a Study of Latin Word Order*. Amsterdam: Benjamins.

Parry, M. (1930) 'Studies in the epic technique of oral verse-making. I. Homer and Homeric style'. *Harvard Studies in Classical Philology* XLI, pp. 73-147 [reprinted in A. Parry (ed.) (1971) *The making of Homeric verse. The collected papers of Milman Parry*, New York-Oxford: Oxford University Press, pp. 266-324].

Ramat, P. (2005) I costrutti assoluti nelle lingue indoeuropee, in P. Ramat, *Pagine linguistiche. Scritti di linguistica storica e tipologica*. Roma-Bari: Laterza.

Schiffrin, D. (1987) *Discourse markers*. Cambridge: Cambridge University Press.

– (1994) 'Making a list'. *Discourse processes* XVII, pp. 377-406.

Üding, G. and B. Steinbrink (1994)[3] *Grundriß der Rhetorik. Geschichte, Technik, Methode*. Stuttgart: Kohlhammer.

Viti, C. (forthcoming) The use of the dual number in Homeric Greek, in *Akten der 13. Fachtagung der Indogermanischen Gesellschaft* (Salzburg, 21th-27th September 2008).

PARTE IV

LA DIMENSIONE MORFOSINTATTICA

I focalizzatori in un corpus di parlato spontaneo (C-ORAL-ROM Italia)

Daniela Acciardi
LABLITA - Università degli Studi di Firenze
daniela.acciardi@gmail.com

Abstract

*The ongoing research has identified that in an Italian spoken corpus some connectors (*perché, quindi, ma, dove, che…*) can operate as a focus markers. The communication will explain in detail the peculiar features of this specific function.*
In fact, connectors can develop many different functions (illocutory functions, Discourse Markers, conjunction, focusing) and the identification of the different ones in a spoken corpus is possible only through listening, analysis and instrumental verification of their features in the utterance. Instrumental tests show a conjunction preserves the value of syntactic and semantic subordination between the clause in the same Information Unit. Otherwise, a focus marker serves only as the phonetic support of the prosodic unit and neutralize its semantic and syntactic value.

1 Introduzione

La presente ricerca intende illustrare le caratteristiche di una specifica funzione connettivale in un corpus di italiano parlato spontaneo (C-ORAL-ROM Italia[1]), la focalizzazione.
In particolare, lo studio corpus-based prevede di differenziare la funzione di focalizzazione dalle altre funzioni connettivali, evidenziarne le caratteristiche fonetiche, prosodiche, distributive, semantiche e sintattiche. Inoltre, l'analisi delle occorrenze del connettivo polivalente ad alta frequenza *perché*, permetterà di rilevare l'incidenza della funzione di focalizzazione.

[1] C-ORAL-ROM Italia è il sottocorpus italiano di C-ORAL-ROM, corpus romanzo (italiano, francese, spagnolo e portoghese) http://lablita.dit.unifi.it/coralrom. Cfr. Cresti and Moneglia (2005).

1.1 Il focus

Il *focus* è da intendere semanticamente come il punto focale dell'informazione, sistematicamente segnalato da caratteri accentuali e da un nucleus intonativo.
In letteratura il *focus* non è stato definito in modo unitario, per cui abbiamo diverse caratterizzazioni di tale concetto: Chomsky (1971) ha definito il *focus* come quell'elemento della frase in cui avviene il cambiamento dell'informazione, contrapponendolo così alla *presupposizione*, che rappresenta il punto in cui il cambiamento informativo deve essere ancorato. In Chomsky and Halle (1968) il termine Focus è utilizzato per riferirsi alle parole o alle espressioni che sono prosodicamente o sintatticamente prominenti, perché introducono informazioni "nuove". Hallyday (1967) afferma che il Focus determina quale parte della frase contribuisce ad informazioni testualmente e situazionalmente non derivabili. Sgall, Hajičová and Benešová (1973) ha definito il *focus* come il segmento della frase che non presenta alcun legame contestuale.
In anni recenti tradizioni di ricerca di origine diversa si sono interessate al *focus*. Avesani and Vayra (2004) notano che esiste una confusione concettuale fra due diverse interpretazioni del termine *focus*. Una categoria interpretativa relazionale del focus lo indica come ciò che è predicato riguardo al Topic. Una seconda interpretazione vede il Focus dal punto di vista referenziale, cioè come

> l'entità sulla quale il parlante vuole richiamare l'attenzione dell'interlocutore, evocando spesso un contrasto con altri referenti attivati dal contesto, membri potenziali di un insieme chiuso, che potrebbero occupare la stessa posizione nella frase. (Avesani and Vayra, 2004: 3)

Frascarelli (2004: CD-ROM) nota però che

> l'intonazione individua e caratterizza il *Focus* come *fenomeno unitario* dal punto di vista *strutturale*, ma ne *distingue* le sottocategorie e le rende immediatamente interpretabili mediante specifiche *caratterizzazioni tonali*. In altre parole, il *Focus* – qualsiasi sia la sua estensione e il suo significato logico-semantico – è individuato dall'intonazione mediante un tono di confine e sono gli eventi tonali ad operare le necessarie diversificazioni, evitando qualsiasi ambiguità interpretativa.

Beaver and Brady (2008) sostengono che il Focus influisce direttamente sulla semantica di una frase. Modi diversi di realizzare una frase possono distinguere un'interpretazione da altre possibili interpretazioni della stessa frase. L'interpretazione è veicolata attraverso il modo in cui le parole sono correlate le une alle altre. Affermano inoltre che il Focus riguarda anche la fonologia ed è veicolato attraverso informazioni soprasegmentali quali il ritmo e l'intonazione e, in alcuni casi, può essere introdotto da *marcatori di Focus*.

La presente trattazione si occuperà nello specifico delle espressioni che svolgono la funzione di introdurre il Focus, inteso come un picco saliente nell'andamento della frequenza fondamentale (F_0) dell'unità tonale di riferimento, cui corrisponde il punto focale dell'informazione entro tale unità d'informazione[2].

1.2 Livelli funzionali della lingua parlata[3]

[2] Si noti che questo è un aspetto che viene riconosciuto fin da subito dagli studiosi indipendentemente dal loro specifico approccio teorico, ad esempio Chomsky (1971) «the focus is the phrase containing the intonation center» e Jackendoff (1972): «the Focus Marker contains a feature marking the pitch contour». Cfr. Selkirk (1995) e Schwarzschild (1999).

[3] I livelli funzionali della lingua fanno riferimento alla Teoria degli atti linguistici di Austin, secondo il quale l'atto linguistico è l'esecuzione simultanea di più atti: (a) Atto locutivo, che comprende la fonazione, la morfologia e la semantica di una lingua; (b) Atto illocutivo, che è l'atto che si compie nel momento in cui si dice qualcosa; ogni enunciato ha una forza illocutiva, ed è questo che fa capire agli interlocutori se la frase pronunciata è una domanda, un ordine, un'asserzione, eccetera; (c) Atto perlocutivo, che s'identifica negli scopi e nelle conseguenze che ha l'atto illocutivo. Cfr. Austin (1962) e John Searle (1969). La schematizzazione presentata fa riferimento alla Teoria della lingua in atto e articolazione informativa (Cresti, 2000), che considera l'Enunciato come l'unità di riferimento della lingua parlata, la cui definizione recita: «enunciato viene definito ogni espressione linguistica interpretabile pragmaticamente, legata: a) ad una condizione semantica di piena significanza dell'espressione in questione (parole lessicale vs morfema); b) alla sua realizzazione intonata secondo un pattern melodico di valore illocutivo. […] Costituisce il corrispettivo linguistico di un atto ed è prosodicamente identificabile nel continuum parlato (criterio illocutivo e articolazione informativa). In tal modo l'enunciato risulta definito dal carattere dell'autonomia, ovvero dell'interpretabilità pragmatica, e l'indice linguistico necessario alla sua realizzazione è l'intonazione» (Cresti, 2005: 251).

Lo studio dei connettivi nella lingua parlata ha tenuto conto dei livelli funzionali realizzati nella lingua parlata:

Livello locutivo
Livello dell'articolazione informativa
Livello illocutivo

Esiste, infatti, un livello locutivo che riguarda le relazioni sviluppate fra le diverse espressioni all'interno dell'Unità Informativa[4] (UI). Il livello dell'articolazione informativa riguarda, invece, i rapporti fra le diverse UI all'interno dell'enunciato. Il livello illocutivo, infine, si riferisce alla forza illocutiva[5] di un enunciato.
Il livello illocutivo e quello dell'articolazione informativa del parlato sono legati all'aspetto sociale e interattivo della lingua. Il livello locutivo, invece, comprende la fonazione, la morfologia e la semantica di una lingua, assumendo un aspetto semantico-modale e un altro sintattico-morfologico.

[4] Nell'ambito della Teoria della lingua in atto (Cresti, 2000), ambito nel quale la ricerca si struttura, l'enunciato è realizzato attraverso un pattern informativo che può essere articolato in più Unità Informative, dedicate ciascuna a funzioni informative diverse. Fra le diverse unità informative, quella primaria è il *Comment* (nella trascrizione degli esempi è segnalato con la sigla in apice ^{COM}) che rappresenta la UI necessaria e sufficiente per la realizzazione e l'interpretabilità di un enunciato poiché realizza la forza illocutiva dell'enunciato. Le altre unità informative sono: il *Topic* (segnalato nella trascrizione con la sigla in apice ^{TOP}), l'*Appendice* (segnalato nella trascrizione con la sigla in apice ^{APC} se appendice di Comment, ^{APT} se di Topic), l'*Inciso* (segnalato nella trascrizione con la sigla in apice ^{PAR}), l'*Introduttore locutivo* (segnalato nella trascrizione con la sigla in apice ^{INT}). All'interno dell'enunciato possiamo trovare anche unità che non hanno una funzione informativa vera e propria, ma riguardano la funzione interpersonale della lingua: gli *Ausili Dialogici*. Anch'essi possono essere di diversi tipi: *Incipit* (segnalato nella trascrizione con la sigla in apice ^{INP}), *Fatici* (segnalato nella trascrizione con la sigla in apice ^{PHA}), *allocutivi* (segnalato nella trascrizione con la sigla in apice ^{ALL}), *espressivi* (segnalato nella trascrizione con la sigla in apice ^{EXP}), *conativi* (segnalato nella trascrizione con la sigla in apice ^{CNT}) e *connettori testuali* (segnalato nella trascrizione con la sigla in apice ^{TXC}). Cfr. Cresti (2000) e Cresti and Moneglia (in press).

[5] Ogni enunciato ha una forza illocutiva realizzata nel Comment attraverso l'intonazione della voce, interpretabile dagli interlocutori poiché determinata convenzionalmente nella lingua; gli interlocutori possono in questo modo interpretare se l'enunciato sia una domanda, un ordine, un'asserzione ecc. Cfr. Cresti (2000).

I connettivi possono essere impiegati entro diversi livelli funzionali e assumono valori differenti a seconda del livello in cui sono operanti.
Un connettivo, dunque, può avere funzione illocutiva se è portatore del focus informativo dell'enunciato; in questi casi, si trova all'interno dell'UI di Comment, esaurendo spesso l'intero enunciato.
Nei casi in cui un connettivo svolge funzioni a livello dell'articolazione informativa, svolge di norma funzioni di ausilio dialogico, soprattutto di Incipit, di Fatico e di Connettore Testuale. Un connettivo in funzione di UI Dialogica non ha nessi con il contenuto locutivo dell'enunciato vero e proprio, poiché, quando un connettivo assolve le funzioni di Incipit, sancisce la presa di turno o la conferma nel corso dello svolgimento del turno stesso. Quando funge da Fatico, sollecita percettivamente l'interlocutore per assicurare l'apertura del canale di comunicazione[6]. Il Connettore Testuale, infine, segnala all'interlocutore la ripresa della funzione precedentemente messa in atto per completarla, manifesta cioè che è in atto un processo di composizione del testo e richiama anaforicamente il contenuto di quanto è stato enunciato in precedenza[7].
Il livello maggiormente interessato dalle funzioni svolte dai connettivi nella lingua è il livello locutivo. È infatti a questo livello cui fanno capo gli usi sia sintattici subordinanti che avverbiali dei connettivi.

1.3 Connettivi ad alta frequenza nel corpus

La ricerca comprende un primo livello di analisi dei dati quantitativi che sono stati messi a confronto con quelli dello scritto letterario. Sono stati rilevati a questo scopo i connettivi più diffusi nel corpus di parlato spontaneo C-ORAL-ROM Italia, un corpus di Italiano spontaneo parlato contenente 310.969 parole trascritte, oltre 36 ore di registrazione, 204 sessioni e 451 parlanti. I dati riguardanti il numero di occorrenze dei connettivi più importanti rilevati in C-ORAL-ROM, sono stati messi a confronto con i dati relativi a GRITTEXT[8], corpus

[6] Per un approfondimento sulle caratteristiche degli ausili dialogici si veda Frosali (2005).
[7] Cfr. Cresti (in press), Scarano (2007) e Acciardi (in press b).
[8] Il corpus di italiano letterario GRITTEXT è stato raccolto e archiviato presso il laboratorio di linguistica italiana del Dipartimento di Italianistica della Facoltà di Lettere e filosofia dell'Università degli studi di Firenze LABLITA. http://lablita.dit.unifi.it/grittext/php/index.php.

di scritto letterario che comprende 234.245 parole, 275 campioni di 171 opere scritte dal 1950 ad oggi, di 77 autori diversi. I due corpora sono confrontabili per numero di parole totali[9] e per grado di rappresentatività delle due varietà diamesiche dell'italiano.

Lemma	C-ORAL-ROM	GRITTEXT
Che	7693	5451
E	6864	7836
Ma	2227	928
Perché	2124	487
Anche	1907	586
Se	1632	839
Come	1511	1365
Poi	1468	373
Cioè	1162	47
Quindi	1039	18
Allora	1012	162
Però	1001	87
O	930	808
Quando	624	403
Così	567	365
Dopo	624	282
Invece	372	139
Comunque	332	33
Dove	291	298

Tabella 1

Il confronto riportato nella tab. 1 mette in evidenza che nel parlato i connettivi (congiunzioni coordinanti, subordinanti e avverbi) ad alta frequenza sono sistematicamente più numerosi che nello scritto.
Come abbiamo notato in precedenza, anche se nel parlato esistono usi connessi alle funzioni illocutive e di supporto dialogico non riscontrabili nello scritto, il grande scarto nel numero delle occorrenze fra parlato e scritto, come illustreremo ampiamente nei paragrafi successivi, non può essere spiegato solo da tali usi.

[9] Il confronto ha tenuto conto della differenza di quasi il 25% fra il numero di parole totali dei due corpora.

1.4 Le funzioni connettivali nel parlato

La letteratura dedicata fornisce diverse definizioni della funzione connettivale. Dressler (1974: 103) afferma che

> i connettivi ripartiscono gli insiemi di frasi, in quanto possono separare le relazioni meno strette da quelle più strette. Le congiunzioni e gli altri connettori sono solamente l'espressione esterna delle relazioni semantiche dell'associazione di frasi, sono elementi facoltativi in parte ridondanti e in parte solo puntualizzanti.

La definizione è stata ripresa e approfondita da Halliday and Hasan (1976: 227), che affermano inoltre che «i connettivi hanno funzione di collegamento fra gli elementi linguistici dello stesso testo che non sono messi in relazione da altri mezzi strutturali». Più di recente, Samardzic (1995: 67) puntualizza:

> i connettivi svolgono la funzione di raccordo tra le varie parti del discorso contribuendo alla sua pianificazione sintattica. Oltre alla funzione di stabilire i rapporti sintattici all'interno di una proposizione o di un periodo, i connettivi possono articolare intere porzioni del testo[10].

In base alle definizioni sopra riportate, bisogna considerare tre diversi tipi di funzioni connettivali: la congiunzione frasale, il focalizzatore e il connettore testuale. I connettivi con funzione di congiunzione frasale stabiliscono i rapporti sintattici fra costituenti all'interno della stessa Unità Informativa (UI). I focalizzatori, come vedremo nei paragrafi successivi, sono elementi facoltativi inseriti nella UI e non hanno funzione subordinante. I connettori testuali, infine, servono ad articolare intere porzioni di testo, creando rapporti paratattici fra le UI interessate. Possiamo affermare, dunque, che queste tre funzioni conservano un valore connettivale, benché facciano riferimento a livelli della lingua differenti.

[10] Cfr. Bazzanella (1985), Ferrari (1995; 2008), Sabatini and Coletti (2005) e Beretta (1984).

1.5 I focalizzatori

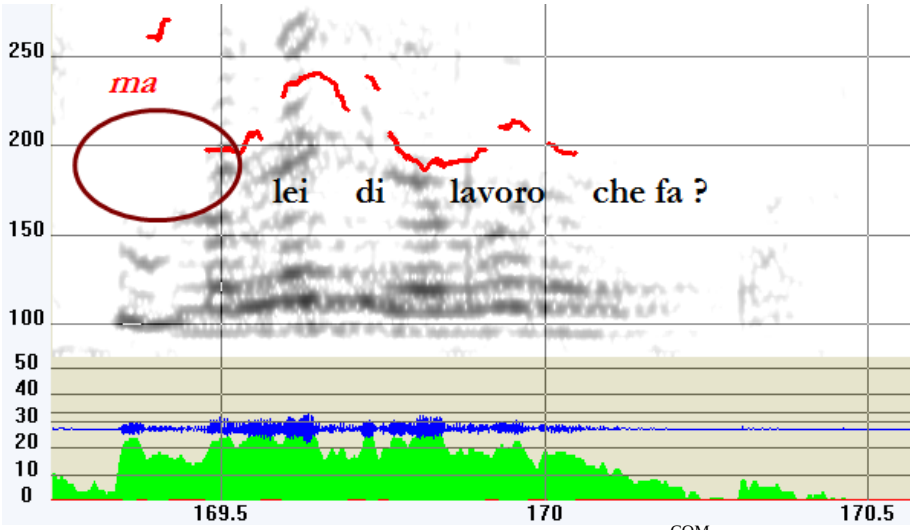
Carla Bazzanella (1995: 247) definisce i focalizzatori come quei connettivi[11] che «possono indirizzare o regolare l'elaborazione dell'informazione a livello cognitivo. […] Possono inoltre sottolineare i punti focali del discorso».
Nella lingua parlata, dato che un focalizzatore svolge la funzione di introdurre il Focus informativo di una UI, è necessario che occorra in UI di tipo testuale che siano portatrici di Focus, cioè il Topic oppure il Comment[12]. Il focalizzatore, attraverso una realizzazione fonetica con intonazione bassa o di preparazione al movimento successivo, serve a far meglio apprezzare la curva intonativa del Focus, sottolineando o esplicitando il valore semantico complessivo dell'enunciato in cui occorre.
Possono svolgere la funzione di focalizzazione tutti i connettivi in generale, ma nel corpus di riferimento sono soprattutto le congiunzioni coordinanti e subordinati e gli avverbi interrogativi a assolvere tale funzione, come ad esempio *ma*, *che*, *quindi*, *perché*, *come*, *dove*, *quando*, *anche*, *allora*.

1.6 Aspetti distributivi, fonetici e prosodici

Un focalizzatore si trova sempre all'inizio della UI testuale che introduce ed è seguito almeno da un'altra espressione. In alcuni casi due o più connettivi brevi possono affiancarsi e svolgere complessivamente la funzione di focalizzazione che prepara al focus della UI. Le espressioni che seguono il focalizzatore sono ben messe in rilievo; infatti, anche quando il focalizzatore è coarticolato ad esse, la fine del connettivo è sempre ben percepibile.
Un connettivo che svolge questa funzione è pronunciato con una velocità più alta della velocità media dell'unità in cui occorre e possiede una realizzazione fonetica debole e ridotta. Anche la sua intensità è più bassa rispetto alle espressioni che seguono.

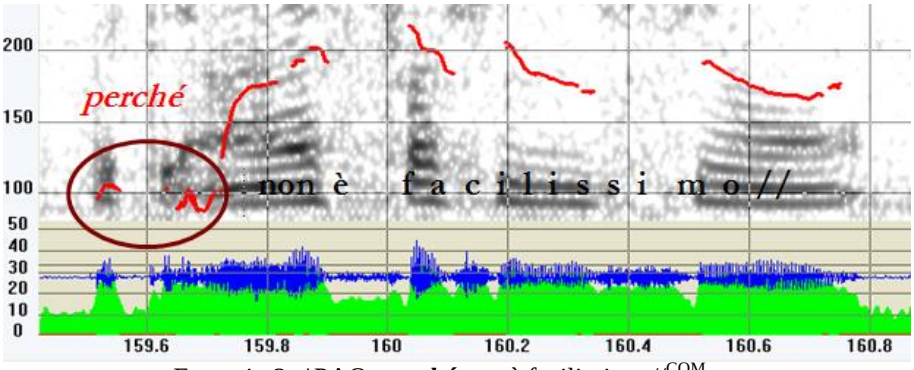
[11] Bazzanella usa il termine *segnali discorsivi* per riferirsi ad un'insieme di funzioni più ampia di quella solitamente associata ai connettivi, ma nella nostra trattazione tutte le funzioni associate ai focalizzatori rientrano in quelle connettivali.
[12] Cfr. Cresti (2000).

Dal punto di vista prosodico, si presenta associato a valori della Frequenza Fondamentale (F_0) più bassi rispetto alla F_0 media delle altre espressioni della UI, come nell'esempio seguente.



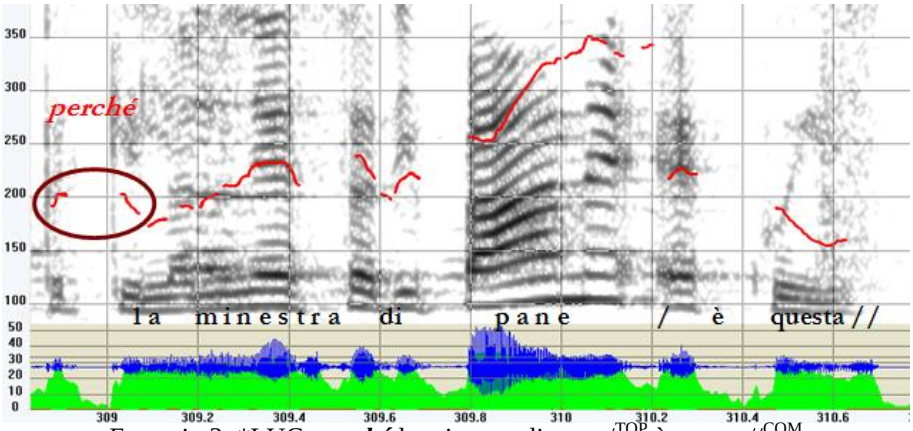
Esempio 1: *CLA: ***ma*** lei di lavoro che fa ?[COM]

Dallo spettrogramma è evidente che il focalizzatore *ma* non è foneticamente pienamente realizzato, ha un'intensità ed una F_0 inferiore alla media del resto dell'enunciato.
Molto spesso il focalizzatore è subito adiacente al focus dell'enunciato (es. 2), ma, in alcuni casi, esiste una sorta di preparazione al focus che inizia appena dopo il focalizzatore (es. 3).



Esempio 2: *PAO: ***perché*** non è facilissimo //[COM]

Nell'esempio 2, il focus è evidentemente immediatamente adiacente al focalizzatore *perché*.



Esempio 3: *LUC: ***perché*** la minestra di pane /[TOP] è questa //[COM]

Nell'esempio 3 il focus informativo è sulla parola *pane*, ma la curva che dà inizio al focus comincia appena dopo il focalizzatore *perché*.
Di fatto un focalizzatore si comporta da appoggio fonetico per ciò che segue.

1.7 Caratteristiche sintattiche e semantiche

Un focalizzatore non svolge alcun tipo evidente di funzione sintattica subordinante o avverbiale e non crea né legami né strutture gerarchiche fra le espressioni che seguono il connettivo e quelle che lo precedono, come è possibile osservare nell'esempio seguente:

Esempio 4: *FAB: se c'è dell'acqua /[TOP] se c'è delle &partice [/1][SCA] delle bollicine d'aria /[TOP] anche invisibili /[APT] l'oggetto esplode /[COM] insomma //[PHA] ***quindi*** deve essere completamente secco //[COM]

Fra la UI introdotta dal focalizzatore *quindi* e la UI precedente non esiste alcun rapporto gerarchico; si tratta infatti di due Comment completamente indipendenti.
All'ascolto degli enunciati, risulta evidente che anche il valore semantico di norma associato al connettivo è fortemente indebolito. Il nesso fra le espressioni che precedono il focalizzatore e quelle che lo seguono, viene realizzato completamente dall'intonazione. La forza

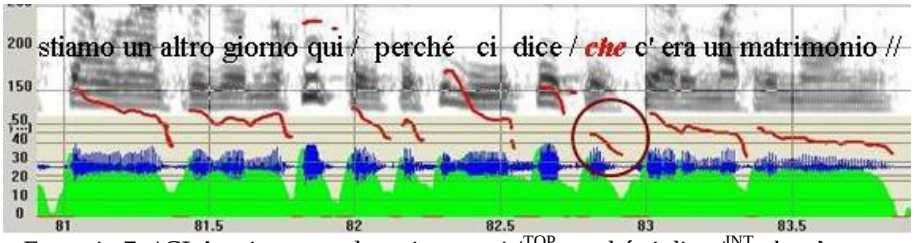
illocutiva complessiva dell'enunciato è coerente con i valori semantici normalmente associati al connettivo scelto[13]. Il focalizzatore potrebbe essere considerato una marca morfologica dell'illocuzione del Comment che introduce.

Esempio 5: *GAB: poi /^{PHA} ***dove*** a Pubo ?^{COM}

Nell'esempio 5 il connettivo *dove* infatti ha valore interrogativo, poiché introduce un Comment di domanda, ma perde la sua funzione avverbiale poiché la domanda è di tipo totale, non parziale. Il parlante non chiede in quale posto di "Pubo", ma chiede se un tale evento è accaduto in quel posto. La domanda totale è quindi "a Pubo?" e il connettivo *dove* serve solo da appoggio per formulare la domanda.

Esempio 6: *ROS: a te te lo fo volentieri /^{COB} ***perché*** tu sei un tipino //^{COM}

Nell'esempio 6, invece, il connettivo *perché* esplicita il valore del nesso semantico esistente fra le unità informative ad esso adiacenti, il secondo Comment infatti ha illocuzione esplicativa.
Bisogna osservare comunque che il connettivo non crea rapporti di subordinazione fra le UI poiché i Comment legati occupano lo stesso piano gerarchico.



Esempio 7: *CLA: stiamo un altro giorno qui /^{TOP} perché ci dice /^{INT} *che* c'era un matrimonio //$^{COM\_r}$

L'esempio 7 mostra un esempio interessante. La costruzione della frase dal punto di vista sintattico prevede che il connettivo *che*

[13] In questo caso parliamo di forza illocutiva dal punto di vista percettivo, senza far riferimento a studi sulle forme intonative della F0 del Comment, ma solo alla sua interpretazione da parte dell'interlocutore o dell'ascoltatore, che sono naturalmente capaci di interpretare correttamente l'aspetto convenzionale dell'atto linguistico effettuato da un parlante.

introduca un discorso indiretto retto dal verbo *dire*. L'ascolto dell'enunciato, invece mostra che il connettivo non svolge alcuna funzione di subordinazione, ma serve da focalizzatore per un discorso diretto riportato.
Gli enunciati interessati sono stati sottoposti ad un procedimento sperimentale di soppressione del segmento sonoro corrispondente al connettivo con funzione di focalizzatore, secondo un sistema spesso utlizzato in LABLITA[14] per l'analisi del valore informativo delle diverse espressioni. Tale soppressione ha offerto un'importante conferma, mostrando che l'interpretazione dell'enunciato non risulta modificata. Invece, nei casi in cui il connettivo assolve a piene funzioni semantiche e sintattiche, gli esperimenti di soppressione sonora generano enunciati agrammaticali, non interpretabili, oppure danno luogo ad interpretazioni dell'enunciato incongruenti con quelle dell'enunciato originale; la struttura gerarchica dell'enunciato originario non è recuperabile.

1.8 Funzioni svolte da *perché*

Al fine di verificare l'incidenza della funzione di focalizzazione nel parlato spontaneo, è stata effettuata un'analisi quantitativa delle funzioni svolte da *perché* all'interno della sezione family/private di C-ORAL-ROM. Esso è infatti uno dei connettivi di maggiore interesse, per la frequenza d'uso e l'intensa diffusione ad ogni livello della lingua.
Nel sottocorpus di riferimento, anche il connettivo *perché*, come gli altri connettivi[15], può essere utilizzato nei diversi livelli funzionali della lingua parlata.

1.9 Funzioni svolte da *perché* a livello illocutivo

Per quanto riguarda il livello illocutivo, sono state considerate tutte le occorrenze di *perché* nelle quali il connettivo è portatore della forza illocutiva dell'intero enunciato.

[14] Laboratorio di Linguistica Italiana del Dipartimento di Italianistica dell'Università degli Studi di Firenze (http://lablita.dit.unifi.it).
[15] Si veda sopra paragrafo 0.

Le funzioni illocutive di *perché* rilevate nel corpus sono solo di due tipi: domande e espressioni di ovvietà[16].
In questi casi, *perché* è portatore diretto della forza illocutiva[17] e la forma prosodica, e in particolare il movimento della frequenza fondamentale (F_0), è quello tipico della domanda o dell'espressione di ovvietà. Come negli esempi seguenti:

Esempio 8: Domanda → *EST: ***perché*** ?^{COM}

Esempio 9: Espressione di ovvietà → *ELA: bah /^{INP} io son rimasta colpita //^{COM} ***perché*** …^{COM}

In questi casi *perché* si trova all'interno della UI di Comment, spesso esaurisce l'intero enunciato.

1.10 Funzioni svolte da *perché* a livello dell'articolazione informativa

Il connettivo *perché*, a livello dell'articolazione informativa dell'enunciato, può svolgere diverse funzioni di Ausilio Dialogico, ma sono state trovate anche occorrenze di *perché* che svolgono funzioni di Introduttore Locutivo[18].
L'uso di *perché* come Introduttore Locutivo riguarda sempre il livello dell'articolazione informativa, ma è l'unico caso che va riferito ad un'unità informativa di tipo testuale. In questi casi, *perché* introduce liste di cause (es. 14), oppure introduce spiegazioni complesse articolate in diverse fasi (es. 15).

Esempio 14: *FRA: insomma /^{INP} no /^{PHA} cioè /^{PHA} son stanca /^{COM} ***perché*** /^{INT} vabbé /^{PHA} un po' /^{TOP} anche l'agitazione /^{CMM} il lavoro /^{SCA} in Nazionale /^{CMM} tutti questi lavori /^{SCA} comunque /^{SCA} nuovi //^{CMM}

[16] L'espressione di ovvietà consiste nella presentazione linguistica di un evento le cui conseguenze siano note; essa risulta sospesa dal punto di vista della locuzione, ma non da quello del compimento illocutivo, che in tal modo manifesta appunto l'ovvietà (Cresti and Firenzuoli, 2001).
[17] Per approfondimenti sulle forme di valore illocutivo si veda Firenzuoli (2003).
[18] Un introduttore locutivo svolge la funzione di segnalare al destinatario che l'enunciato che segue ha un valore meta-illocutorio, e perciò la sua forza illocutiva non ha valore nel mondo; cfr Cresti (2000).

Esempio 15: *LID: sì /[INP] ***perché*** /[INT] buttan l'ancora in un posto /[CMM] quello co' le bombole va giù /[CMM] e sa' 'ndo vann' a finire ?[CMM]

Le occorrenze in cui il connettivo svolge funzioni di Ausilio Dialogico prevedono casi in cui svolge funzioni di Incipit, di Fatico o di Connettore Testuale, ma, in alcuni casi, anche di Conativo[19].
Quando *perché* ha la funzione di Incipit si trova ad inizio enunciato ed è realizzato tramite un'unità prosodica dedicata, senza relazioni con il resto dell'enunciato, né dal punto di vista sintattico, né semantico (es. 10).

Esempio 10: *ZIA: ***perché*** /[INP] Dio ce ne guardi /[TOP] lu 'mmazzavono //[COM]

Nell'esempio 10, la F_0 in corrispondenza del connettivo è quella tipica degli Incipit, cioè breve, con un attacco alto e in discesa.
Nei casi in cui *perché* assume, invece, la funzione di Fatico, può avere una distribuzione libera all'interno dell'enunciato e il connettivo è segnalato da un'unità prosodica dedicata (es. 11).

Esempio 11: *NIC: ***perché*** /[PHA] non so se li conosci //[COM]

Nell'esempio 11, i valori della F_0 in corrispondenza di *perché* sono medio-bassi, e l'unità prosodica ha una durata breve e scarsa realizzazione fonetica.
Sono stati inoltre trovate alcune occorrenze in cui *perché* fa parte di locuzioni fisse (*perché sa*, *perché guarda* ecc.) che svolgono complessivamente la funzione di Conativo (es. 13).

Esempio 13: *LID: ***perché*** capisci /[CNT] la macchina /[TOP] l'ha sostituito tante cose //[COM]

Perché con funzione di Connettore Testuale è realizzato tramite un'unità prosodica dedicata. Esso per lo più occorre all'interno di Stanze[20] ed è utilizzato come connessione fra due Comment legati, la

[19] I Conativi spingono l'interlocutore a reagire in maniera adeguata durante lo scambio dialogico o interrompono uno scorretto comportamento dello stesso.; cfr. Cresti (2000).
[20] La Stanza è un'entità linguistica che non corrisponde alla realizzazione di un atto linguistico con una propria forza illocutoria, ma il suo obiettivo primario è la realizzazione di un testo orale, seppure sia sempre realizzata all'interno di una

cui relazione è di tipo paratattico. Perde, quindi, completamente il valore di congiunzione subordinante, ma sembra conservare il suo valore semantico (es. 12).

Esempio 12: *AIA: scusa eh /[CON] io non voglio / replicare /[COB] ***perché*** /[TXC] non è stata una comunicazione /[COB] la tua //[APC]

Quando *perché* svolge la funzione di connettore testuale (TXC), assume caratteristiche fonetiche e prosodiche peculiari: ha durata lunga (compresa fra 500 e 1000 ms), la forma del profilo prosodico è piatta o modulata, e la F_0 ha valori più bassi dei valori della F_0 media delle UI adiacenti.
In alcuni casi, *perché* con funzione di connettore testuale può trovarsi anche ad inizio enunciato, e, in tal caso, sottolinea un legame a livello testuale fra due enunciati diversi ed indipendenti.

1.11 Funzioni svolte da *perché* a livello locutivo

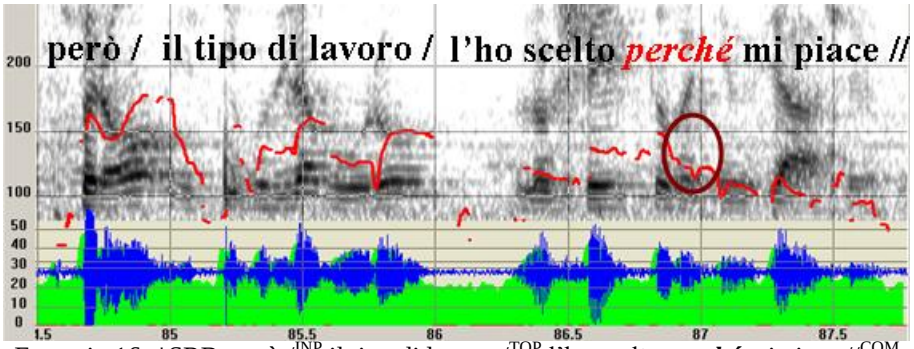
A livello locutivo *perché* può svolgere una funzione di congiunzione frasale oppure quella di focalizzatore. La congiunzione frasale fa riferimento all'aspetto morfo-sintattico della lingua, mentre il focalizzatore ha valore a livello semantico-modale.

1.11.1 *Perché* congiunzione frasale

Perché con funzione di congiunzione frasale serve a legare i costituenti sintattici che appartengono alla stessa UI. Data la peculiarità dell'uso, *perché* con valore di congiunzione si trovano esclusivamente in UI di tipo testuale, preferibilmente all'interno di Comment o di Topic. In questi casi, *perché* lega i costituenti sintattici che appartengono alla stessa UI, conservando la sua funzione avverbiale o subordinante, a seconda dei casi. La F_0 in corrispondenza del connettivo, la velocità e l'intensità, con cui è realizzato il connettivo sono pienamente integrate entro il profilo prosodico della UI di cui fa parte, ovvero il connettivo risulta foneticamente linearizzato, come nell'esempio 16:

---

relazione interattiva con uno o più interlocutori. Cfr. Cresti (2000), Cresti (in press) e Cresti and Moneglia (in press).



Esempio 16: *SRB: però /[INP] il tipo di lavoro /[TOP] l'ho scelto ***perché*** mi piace //[COM]

Dunque, in questi esempi *perché* conserva pieno valore semantico e sintattico subordinante e/o avverbiale, poiché partecipa della composizione del testo dell'enunciato. Nei casi in cui *perché* svolge la funzione di congiunzione frasale, sono realizzati tutti i valori semantici che le grammatiche associano al connettivo polivalente *perché*[21]: esplicativi, causali, finali, consecutivi, completivi, interrogativi.
Come negli esempi seguenti:

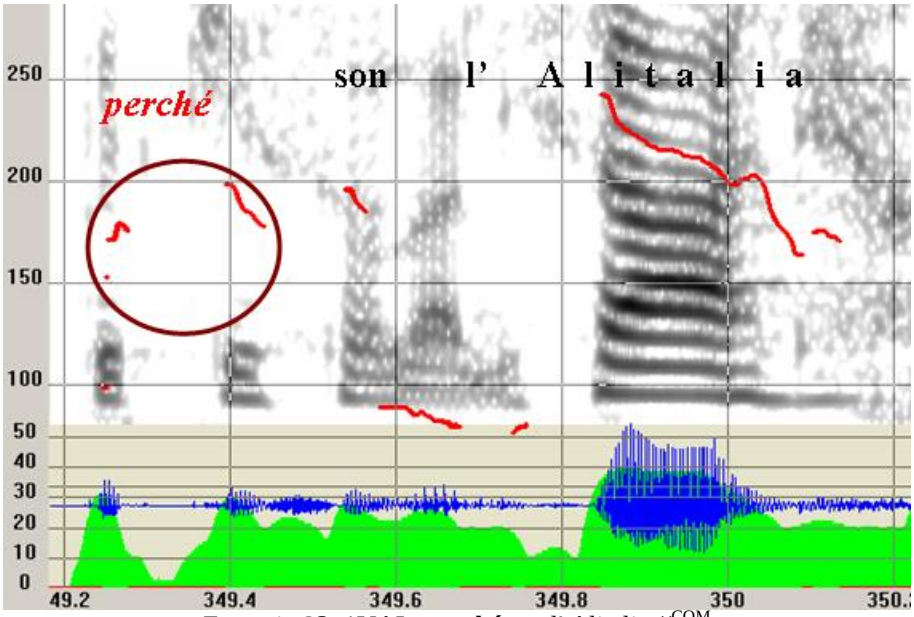
Esempio 17: Esplicativa → *LUC: non tanto piccini ***perché*** poi deve /[COM-i] eh /[PHA] bollire tanto /[SCA] tantissimo //[COM]
Esempio 18: Causale → *SCA: ma se le persone vengono ***perché*** vivono una persecuzione /[TOP] beh allora /[TOP] le porte devono essere spalancate //[COM]
Esempio 19: Finale → *MAU: con /[SCA] una classe professionale intorno altrettanto efficiente ***perché*** lo sviluppo sia armonico a seguire il mercato //[COM]
Esempio 20: Consecutiva → *ANT: e poi lui è troppo presuntuoso ***perché*** io ci lasci stare //[COM]
Esempio 21: Completiva → *BRA: non c'è assolutamente alcuna ragione ***perché*** i pazienti /[SCA] non abbiano diritto alle stesse cure come gli altri //[COM]
Esempio 22: Interrogativa indiretta→ *VAL: non so ***perché*** l'ho fatta //[COM]

1.11.2 Perché focalizzatore

Infine, a livello locutivo, *perché* può svolgere anche la funzione di focalizzazione, che è preminente per l'uso parlato dei connettivi.

[21] In questo gruppo sono state inserite inoltre le occorrenze di *perché* con valore profrase, sostantivato e le locuzioni focalizzati tipo *è perché* ed *ecco perché*, poiché ne condividono il livello morfo-sintattico e le caratteristiche fonetiche e prosodiche.

Come abbiamo notato in precedenza, un focalizzatore può comparire solamente in UI che siano portatrici di Focus informativo: cioè il Topic ed il Comment. La sua realizzazione prosodica è caratterizzata da F_0 con valori più bassi rispetto alla F_0 media delle altre espressioni della UI e la forma della curva della F_0 è piatta ed ha breve durata (fra 150 e 300 ms), come si può osservare nell'esempio 23.



Esempio 23: *VAL: ***perché*** son l' Alitalia //[COM]

I valori semantici associati a *perché* con funzione di focalizzazione devono essere sempre recuperabili dal contesto, ed è l'illocuzione dell'enunciato a fornire di volta in volta l'interpretazione al connettivo.

Nella maggioranza dei casi, il Comment introdotto ha valore esplicativo, come nell'esempio precedente. Altre volte, invece, il focalizzatore *perché* introduce Comment con un'illocuzione di domanda. In questi casi, la domanda non è di tipo parziale, come ci si attenderebbe, ma totale, come nell'esempio 24.

Esempio 24: *DEM: ***perché*** ti registrava /[COM] ogni tanto ?[APC]

1.12 Incidenza delle funzioni svolte da *perché* nel corpus

Abbiamo osservato che il connettivo polivalente *perché* svolge funzioni diversificate ad ogni livello della lingua. Il grafico seguente mostra la distribuzione del connettivo all'interno dei diversi livelli funzionali della lingua. I dati sono relativi alla sezione family/private di C-ORAL-ROM Italia.

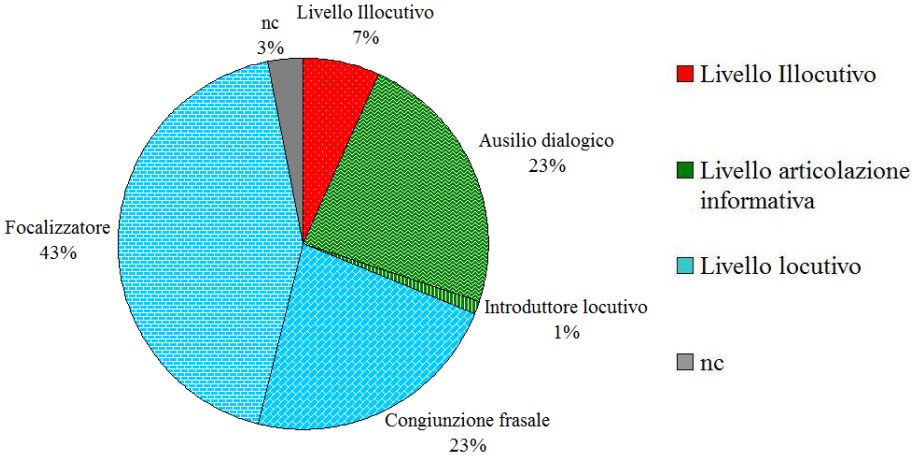


Grafico 1

I dati quantitativi riportati nel grafico 1 mostrano che la funzione di focalizzazione assume un ruolo fondamentale e ricopre la maggior parte delle occorrenze di *perché* nel corpus.

2 Conclusioni

Nel corso della presente trattazione abbiamo identificato tre tipologie diverse di funzioni connettivali, facendo riferimento alle trattazioni classiche e maggiormente condivise sull'argomento. La funzione di focalizzazione, nello schema presentato, rappresenta una particolare tipologia della funzione connettivale. Secondo l'accezione di Dressler, la funzione connettivale è realizzata anche nei casi in cui gli elementi da connettere condividano una relazione semantica indipendentemente dall'operazione svolta dal connettivo. Secondo questa accezione, dunque, la funzione connettivale interessa elementi pleonastici dell'enunciato, il cui compito è quello di indicare una relazione

semantica già presente. Lo studio sul corpus ha permesso di notare che i focalizzatori possono evidenziare oltre che una relazione semantica già in atto, una relazione pragmatica esistente fra diverse UI dell'enunciato; in alcuni casi esplicitano il valore illocutivo del Comment che introducono. Naturalmente, non svolgono funzioni sintattiche avverbiali o subordinanti.
Un connettivo che svolge la funzione del focalizzatore assume caratteristiche fonetiche, prosodiche e distributive ben determinate:

È immediatamente preceduto da break prosodico terminale o non terminale;
Non occorre in unità prosodica dedicata;
Presenta una forma di F0 piatta o leggermente discendente;
La sua F0 è bassa rispetto al resto dell'espressione che lo segue entro l'unità tonale in cui occorre;
Ha una durata breve;
Funge da appoggio fonetico dell'unità tonale in cui occorre;
Aiuta a meglio apprezzare la curva intontiva dell'unità in cui occorre.

L'analisi quantitativa delle funzioni svolte dal connettivo *perché* ha mostrato che la funzione di focalizzazione è molto diffusa. La maggior parte delle occorrenze di *perché* presenti nella sezione family/private di C-ORAL-ROM è infatti interessata da questo uso peculiare, coprendo circa il 43% delle occorrenze di *perché* analizzate (circa 1000).

Riferimenti bibliografici

Acciardi, D. (in press a) *L'uso di perché nel parlato spontaneo (archivio C-ORAL-ROM)*, in Actes du XXVe CILPR.

– (in press b) *Congiunzioni frasali e connettori testuali (C-Oral-Rom e Grittext)*, in A. Ferrari (ed.) *Sintassi storica e sincronica dell'italiano. Subordinazione, coordinazione e giustapposizione. Atti SILFI2008*.

– (in press c) *La subordinazione causale esplicita in un corpus di italiano parlato spontaneo (C-ORAL-ROM Italia)*, in Atti SLI2008.

Austin J.L. (1962) *How to Do Things with Words.* Oxford: Oxford Univerity Press.

Avesani, C. and M. Vayra (2004) *Focus ristretto e Focus contrastivo in italiano*, in F. Albano Leoni, F. Cutugno, M. Pettorino and R.

Savy (eds) *Atti del Convegno "Il parlato italiano", Napoli, 13-15 febbraio 2003*, Napoli: M. D'Auria Editore, CD-ROM, F01, pp. 1-20.

Bazzanella, C. (1985) L'uso dei connettivi nel parlato: alcune proposte, in A. Franchi De Bellis and L.M. Savoia (eds) *Sintassi e morfologia della lingua italiana d'uso. Teorie e applicazioni descrittive. Atti del XVII Congresso Internazionale di Studi. Urbino 11-13 settembre 1983*, Roma: Bulzoni, pp. 83-94.

– (1995) I segnali discorsivi, in L. Renzi, G. Salvi and A. Cardinaletti (eds) *Grande Grammatica di Consultazione*, vol. III, Bologna: il Mulino, pp. 225-257.

Beaver, D.I. and Z.C. Brady (2008) *Sense and Sensitivity: How Focus Determines Meaning*. Malden, MA: Blackwell Publishing.

Berretta, M. (1984) *Connettivi testuali in italiano e pianificazione del discorso*, in L. Coveri (ed.) *Linguistica testuale. Atti del XV congresso internazionale di studi*, Roma: Bulzoni, pp. 237-54.

Chomsky, N. (1971) Deep Structure, Surface Structure and Semantic Interpretation, in D.D. Steinberg and L.A. Jakobovits (eds) *Semantics. An Interdisciplinary Reader in Philosophy, Linguistics and Psychology*, Cambridge: Cambridge University Press, pp. 183-216.

Chomsky, N. and M. Halle (1968) *The sound pattern of English*. New York: Harper and Row.

Cresti, E. (2000) Introduzione, in *Ead.* (ed.), *Corpus di italiano parlato*, Firenze: Accademia della Crusca.

– (2005), *Enunciato e frase: teoria e verifiche empiriche*, in M. Biffi, O. Calabrese and L. Salibra (eds) *Italia Linguistica: discorsi di scritto e di parlato. Scritti in onore di Giovanni Nencioni*, Siena: Prolagon, pp. 249-260.

– (in press) *La parataxe: articulation informationnelle dans le parlé spontané vs juxtaposition syntaxique dans l'écrit littéraire?*, in M. Avanzi and J. Beguelin (eds) *La parataxe*.

Cresti, E. and V. Firenzuoli (2001) 'Illocution et profils intonatifs de l'italien'. *Revue Française de Linguistique Appliquée* 4, 2, pp. 77-98.

Cresti, E. and M. Moneglia (eds) (2005) *C-ORAL-ROM. Integrated reference corpora for spoken romance languages*. Amsterdam: Benjamins.

– (in press) *The annotation of Information Units in spontaneous spoken Italian. Specification of the Textual units (COM TOP APC PAR INT APT SCA).*
Dressler, W.U. (1974) *Introduzione alla linguistica del testo*. Roma: Officina Edizioni.
Ferrari, A. (1995) *Connessioni. Uno studio integrato della subordinazione avverbiale*. Ginevra: Editions Slatkine.
– (2008) *Congiunzioni frasali, congiunzioni testuali e preposizioni: stessa logica, diverso valore semantico-testuale*, in E. Cresti (ed.) *Prospettive nello studio del lessico italiano, Atti SILFI 2006*. Firenze: FUP, pp. 411-6.
Firenzuoli, V. (2003) *Le Forme Intonative di Valore Illocutivo dell'Italiano Parlato: Analisi Sperimentale di un Corpus di Parlato Spontaneo (LABLITA)*. PhD thesis. Dottorato di linguistica, Università degli Studi di Firenze.
Frascarelli, M. (2004) *L'interpretazione del Focus e la portata degli operatori sintattici*, in F. Albano Leoni, F. Cutugno, M. Pettorino and R. Savy (eds) *Il Parlato Italiano. Atti del Convegno Nazionale 13-15 febbraio 2003*, B06, Napoli: M. D'Auria Editore.
Frosali, F. (2005) *Le unità di informazione di ausilio dialogico: valori percentuali, caratteri intonativi, lessicali e morfosintattivi in un corpus di italiano parlato (C-ORAL-ROM)*. Tesi di Laurea, Università degli studi di Firenze.
Halliday, M.A.K. (1967) 'Notes on Transitivity and Theme in English (Part 2)'. *Journal of Linguistics* 3.
Halliday, M.A.K. and R. Hasan (1976) *Cohesion in English*. Londra: Longman.
Jackendoff, R. (1972) *Semantic Interpretation in Generative Grammar*. Cambridge: Cambridge University Press.
Sabatini, F. and V. Coletti (2005) *Il Sabatini Coletti. Dizionario della lingua italiana 2006*. Bologna: Rizzoli.
Samardžić, M. (1995) 'Una proposta per la tassonomia dei connettivi testuali in italiano'. *Italica Belgradensia*, 4, Belgrado, pp. 67-79.
Scarano, A. (2007) *The informational patterning in spoken romance languages*, in G. Banti and L. Mereu (eds) *Structures of information*, Berlin-New-York: Mouton.
Schwarzschild, R. (1999) *'Giveness, AvoiF and other Constraints on the Placement of Accent'*. *Natural Language Semantics* 7, 2, pp. 141-177.

Searle, J. (1969) *Speech Acts: An Essay in the Philosophy of Language*. Cambridge, Eng.: Cambridge University Press.
Selkirk, E. (1995) *Sentence Prosody: Intonation, Stress and Phrasing*, in J.A. Goldsmith (ed.) *The Handbook of Phonological Theory*, London: Basil Blackwell, pp. 550-69.
Sgall, P., E. Hajičová and E. Benešová (1973) *Topic, Focus, and Generative Semantics*. Kronberg: Taunus.

La dislocazione a destra tra italiano comune e variazione regionale

Claudia Crocco
Vakgroep Romaanse talen (andere dan het Frans) - Universiteit Gent
Claudia.Crocco@UGent.be

Abstract

Right dislocation (RD) is usually included in the linguistic features of the so-called neo-standard Italian. This pan-Italian variety is used in relatively informal contexts by speakers with education ranging from average to higher. Although neo-standard linguistic features have been studied from different points of view, their actual geographical diffusion has been not always investigated in depth. By means of a corpus-based analysis the question is addressed if RD is equally represented in different parts of Italy. The analysis has been carried out on the corpus CLIPS which includes task-oriented dialogues collected in fifteen Italian cities. The results have shown that RDs appear seldom in the conversation of speakers form Milan, while they are widespread, among others in the production of the speakers from Florence. The difference between Milan and Florence has been further investigated using the corpus LIP. The comparison between CLIPS and LIP corroborated the results, which are statistically significant: although the RD is present in all the investigated varieties of Italian, its frequency varies significantly in the two cities. This suggests that the pan-Italian character of the neo-standard linguistic features could be re-thought in terms of frequency of occurrence of each feature in different areas. Finally, the deeper linguistic standardization of Milan is invoked to explain the scarce presence of RD in this area.

1 Introduzione

Nei primi studi dedicati alla varietà centrale del repertorio linguistico italiano, il cosiddetto "italiano dell'uso medio" (Sabatini, 1985), "tendenziale" (Mioni, 1983) o "neo-standard" (Berruto, 1987), era individuato un insieme di tratti fonologici, morfosintattici e lessicali panitaliani, estranei allo standard tradizionale e riscontrabili nelle produzioni, soprattutto informali, degli italofoni di media cultura. Questi tratti costituivano il nucleo di una "nuova" varietà, diversa sia dallo standard, sia dalle varietà marcate diastraticamente e diafasicamente come più basse.

La nuova varietà emergeva dall'osservazione della situazione linguistica italiana a cavallo tra gli anni Settanta e Ottanta, in cui una compagine di varietà intermedie veniva definendosi nello spazio linguistico che separa lo standard tradizionale dai dialetti. L'identificazione stessa della nuova varietà media sovraregionale poneva in modo immediato il problema del suo rapporto con le altre varietà di repertorio contigue, caratterizzate in senso diafasico, diamesico o diastratico. Da una parte, ad esempio, l'ingresso di tratti tradizionalmente ritenuti sub-standard, e tipici del parlato in generale, nella varietà centrale del repertorio richiedeva che si analizzasse il rapporto di questa con l'italiano popolare e col parlato; dall'altra, gli studi mostravano l'emergere di tendenze di ri-standardizzazione (Berruto, 1987; Cortelazzo, 2001), riscontrabili nei segni di erosione mostrati dallo standard in seguito all'espansione della nuova varietà. Per approfondire la conoscenza del rapporto tra varietà media e varietà contigue si auspicarono studi sulla frequenza e la distribuzione dei vari tratti in diversi contesti comunicativi (Berruto, 1987; Sabatini, 1985).

Data la complessa trama di relazioni che legano le diverse varietà del repertorio italiano, non sorprende il fatto che, negli studi linguistici degli ultimi decenni, l'interesse per la varietà media si sia intrecciato con quello più generale per il parlato (Della Corte *et al*, 2004; Albano Leoni and Giordano, 2005; De Mauro, 1994; Sobrero and Lavinio, 1991) e con quello per lo sviluppo diacronico della lingua italiana, attraverso i quali si è rilevata la persistenza di tratti linguistici non standard in diverse varietà e fasi dello sviluppo storico della lingua (D'Achille, 1990).

Sono stati approntati inoltre studi su un insieme di fenomeni tradizionalmente esclusi dallo standard, studi che hanno contribuito più o meno direttamente a definire in maggiore dettaglio le caratteristiche della varietà media. In alcuni casi, tali indagini hanno mostrato la relativa rarità di un certo tratto nei testi ascrivibili a tale varietà. Ad esempio, accurati e ripetuti sondaggi sull'uso esteso del pronome *che* per la formazione di clausole relative non standard (Alfonzetti, 2002; Aureli, 2003; 2004; Aureli *et al*, 2005; Bernini, 1991), tratto già presente in fasi antiche dello sviluppo linguistico dell'italiano (Berruto, 1987; Fiorentino, 1999) e ritenuto caratteristico della varietà media, hanno mostrato che si tratta di un uso tipico delle produzioni sub-standard (Alfonzetti, 2002) e non maggioritario nel

parlato definibile come colloquiale (si vedano in particolare Aureli (2003; 2004), in cui sono discussi dati ricavati da un campione di parlato molto ampio e diversificato). La diffusione abbastanza limitata della relativa non standard suggerisce che essa, pur rappresentando una risorsa linguistica importante (Fiorentino, 1999), non sia un tratto non centrale della varietà media, ma che sia invece connotata in senso diastraticamente e/o diafasicamente basso (cfr. Berruto, 1987: 123-25).

Risultati come quelli citati pongono quindi il problema di quale sia l'attuale peso effettivo dei singoli tratti individuati nei primi studi sulla varietà media nel caratterizzare la varietà stessa, ovvero di quali tratti siano centrali e quali marginali per la definizione della varietà.

Poiché la varietà in questione è definita non solo in senso diafasico come *media*, ma anche in senso diatopico come *comune*, o sovraregionale, il suo rapporto con le varietà geografiche è di grande importanza (cfr. bibliografia citata in Radtke, 1992). Trattandosi di una varietà media *comune*, essa include, infatti, per definizione, soltanto tratti presenti su tutto il territorio italiano, differenziandosi quindi dalle varietà regionali proprio per la più ampia diffusione dei suoi tratti caratteristici. Ciò non implica, però, che il rapporto che lega italiano medio comune e varietà regionali sia esclusivamente definibile in termini di reciproca delimitazione (ciò che non è comune è regionale, e viceversa). Nell'interazione tra italiano comune e italiani regionali, ad esempio, potrebbe avere un ruolo importante la frequenza di ciascun tratto dell'italiano comune nelle diverse regioni. La diffusione sovraregionale di un certo fenomeno, cioè, non assicura automaticamente che esso sia ugualmente frequente in ogni parte del territorio: un tratto può essere sì presente in tutte le regioni italiane, ma può anche essere più frequente in alcune e meno in altre.

È questo aspetto del problema della definizione della varietà media che si intende sviluppare in questo lavoro, benché limitatamente alla diffusione geografica di una sola delle costruzioni morfosintattiche normalmente assegnate alla varietà media, cioè la dislocazione a destra (DD). Attraverso l'esame della diffusione di questo tratto ci si propone di contribuire ad una definizione più accurata della varietà media comune rispetto alle possibili variazioni diatopiche.

I dati quantitativi raccolti in questo lavoro sono stati ricavati da uno spoglio estensivo su dialoghi spontanei di quindici varietà regionali

diverse tratti dal corpus CLIPS dialogico e sono stati integrati con un campione di dialoghi del corpus LIP.
Il lavoro è articolato come segue: nel paragrafo §2 è brevemente descritta la DD con ripresa pronominale trattata in queste pagine; nel paragrafo successivo (§3) è presentato il corpus utilizzato; quindi (§4) si dà una descrizione delle caratteristiche generali delle DD rinvenute nei dialoghi (numerosità, tipi sintattici, modalità). Nella sezione §4.2 sono raccolti i dati relativi alla numerosità e alla distribuzione diatopica della DD nelle quindici località presenti nel corpus CLIPS. Come si vedrà, in CLIPS la costruzione appare distribuita geograficamente in modo non omogeneo. Per verificare questo risultato è stato quindi allestito un confronto (§4.3) con il corpus LIP di Firenze e Milano. Poiché il dato ricavato da CLIPS è in linea generale confermato, anche dal punto di vista statistico (§4.3.1), dall'analisi delle sezioni del LIP, in §5 è proposta una possibile pista interpretativa per rendere conto delle differenze osservate nella distribuzione. I risultati sono quindi iscritti nel quadro più generale dello studio della varietà media di italiano (§6).

2 La dislocazione a destra con ripresa clitica

In questo studio sono prese in esame soltanto le DD che presentano una ripresa clitica del costituente dislocato. Tale costruzione è presente in italiano già in fasi molto antiche (D'Achille, 1990) ed ha nell'italiano informale attuale una diffusione abbastanza ampia, benché solitamente ritenuta inferiore rispetto alla corrispondente dislocazione a sinistra (Berruto, 1986; Bozzone Costa, 1991: 151); risultati diversi sono in Rossi (1999) e Messina (2007: 210-12), sul parlato filmico e su quello televisivo rispettivamente, nei quali la DD occorre più frequentemente di quella a sinistra.
Di seguito sono riportati alcuni esempi di DD con ripresa pronominale:

(1) *Lo porti domani il dolce?*
(2) *Lo porto domani il dolce*
(3) *Quanti ne porti di dolci?*
(4) *Ne porto tre di dolci*
(5) *Lo sai che Luca porta il dolce*

Le costruzioni esemplificate in (1)-(5) non rappresentano gli unici tipi di DD possibili in italiano (Benincà, 2001²) nei casi (1), (2) e (5), infatti, sono ammesse dislocazioni simili anche senza la ripresa pronominale (escluse invece per (3) e (4); cfr. Calabrese, 2001²: 577; Cordin, 2001²: 651). In assenza di raddoppiamento clitico, però, la costruzione dislocata (in questo caso definibile come "emarginazione"; Cardinaletti, 2002) non presenta tratti sintattici specifici ed è quindi più difficile da identificare. L'esempio (6), ripreso da Benincà (2001²: 161) illustra un caso in cui è assente la ripresa pronominale e il costituente nominale non segue direttamente il verbo, ma è separato da questo da un avverbio e da una discontinuità prosodica resa graficamente con la virgola:

(6) *Porto domani, il dolce*

Simile alla dislocazione dell'oggetto senza raddoppiamento è la dislocazione del soggetto. Anche in questo caso, si tratta di una costruzione che non può essere agevolmente identificata solo su base sintattica, sia per l'assenza di clitici soggetto in italiano, sia per la difficoltà di distinguere tra soggetto postverbale neutro e soggetto dislocato nei casi in cui il verbo ammetta la posizione postverbale del soggetto come opzione basica (Bonvino, 2004; 2005).
Anche nel caso del soggetto la presenza di discontinuità prosodiche ha un ruolo nell'espressione della dislocazione: come mostrano studi recenti (Bonvino, 2004; 2005) soggetto postverbale dislocato e soggetto postverbale non dislocato sono distinti, oltre che da una serie di caratteristiche sintattiche, semantiche, lineari e di modalità, anche da specifici tratti prosodici.
Nel complesso, la dislocazione con ripresa pronominale ha ricevuto più attenzione da parte degli studiosi rispetto alla variante senza ripresa. I dati presentati in queste pagine vanno quindi a integrare le ricerche svolte in precedenza su questa costruzione (tra gli altri da Berruto, 1986; Bozzone Costa, 1991; D'Achille, 1990; Messina, 2007; Rossi, 1999; 2005; Simone, 1997), con dati empirici raccolti analizzando dialoghi regionali di registro informale.

3 Corpus

Per questo studio sono stati utilizzati dialoghi tratti dal corpus CLIPS (Corpora e Lessici di Italiano Parlato e Scritto; http://www.clips.unina.it; Albano Leoni, 2003) e dal corpus LIP (Lessico di frequenza dell'Italiano Parlato; De Mauro *et al*, 1993). Di seguito sono descritte le caratteristiche rilevanti dei due corpora e sono specificate le dimensioni dei campioni.

3.1 CLIPS

La sezione parlata del corpus CLIPS raccoglie numerosi tipi di parlato, tra cui dialoghi orientati di due diversi tipo ("test delle differenze" e *map task*; Anderson *et al*, 1991). I locutori coinvolti in questi dialoghi non possono vedersi e devono svolgere dei compiti (trovare le differenze tra due vignette o ricostruire un percorso) esclusivamente dialogando.
Per questo lavoro è stata analizzata la porzione trascritta ortograficamente del corpus CLIPS dialogico, che contiene 45 dialoghi *map task* e 45 dialoghi del tipo "test delle differenze" raccolti in 15 città italiane. La scelta delle città presenti in CLIPS è stata fatta sulla base di uno studio sociolinguistico (Sobrero and Tempesta, 2007) che ha individuato località rappresentative dal punto di vista linguistico e demografico. Dal punto di vista della variazione regionale, quindi, CLIPS rappresenta un'immagine bilanciata e fedele della realtà linguistica italiana dell'inizio degli anni Duemila.
Per lo studio sono stati analizzati in totale 90 scambi prodotti da parlanti di età compresa tra i venti e i trent'anni, che frequentano, o hanno frequentato, l'università e conoscono personalmente l'interlocutore con cui dialogano. La durata complessiva dei dialoghi esaminati è di circa 17 ore e mezza; il campione consta di circa 62.000 parole grafiche[1] (cfr. tabella 1).

[1] È opportuno specificare che, nel computo delle parole grafiche, non si è tenuto conto dell'occorrenza di eventuali polirematiche; il dato non è quindi del tutto completo. Si può supporre comunque che in un corpus di dialoghi orientati l'incidenza di tali lessemi sia inferiore rispetto a quanto avviene in un corpus come il LIP, caratterizzato da una grande variabilità di registri e contesti comunicativi, dove le polirematiche rappresentano un insieme cospicuo (1933 lessemi complessi sul totale del LIP; De Mauro and Voghera, 1996: 110).

3.2 LIP

Come è noto, il corpus LIP è diversificato al suo interno sia dal punto di vista diafasico, sia dal punto di vista diatopico (De Mauro *et al*, 1993). Delle quattro località rappresentate nel LIP sono state esaminate, per ragioni che saranno esposte più avanti (cfr. §4.2), soltanto Milano e Firenze.

Città	**Minutaggio**	**Parole grafiche**	**Fluenza (parole/minuto)**
Bari	0.43.03	1979	46,02
Bergamo	1.02.57	2629	42,40
Cagliari	1.15.35	3789	50,52
Catanzaro	0.50.22	1538	30,76
Firenze	1.33.52	5917	63,62
Genova	0.50.14	3113	62,26
Lecce	0.49.26	2440	49,79
Milano	1.15.26	5002	66,69
Napoli	1.10.04	3787	54,1
Palermo	1.09.49	4152	60,17
Parma	1.16.05	5150	67,76
Perugia	1.13.05	3399	46,56
Roma	1.19.30	4604	58,27
Torino	1.48.48	10489	97,12
Venezia	1.15.10	4104	54,72
Totale	**17.33.26**	**62092**	**58,96**

Tabella 1: Il dato sul minutaggio è tratto da http://www.clips.unina.it

Per ciascuna di queste due località sono stati analizzati i testi di tipo A e B e cioè, rispettivamente, gli scambi bidirezionali con presa di parola libera faccia a faccia (cioè conversazioni libere tra interlocutori compresenti) e gli scambi bidirezionali con presa di parola libera non faccia a faccia (per es. telefonate tra privati o tra ascoltatori e conduttori di trasmissioni radiofoniche). La scelta di questi due gruppi di conversazioni si deve alla somiglianza tra i dialoghi LIP di tipo A e B e i dialoghi orientati del CLIPS. Le sezioni esaminate, infatti, sono quelle tendenzialmente più informali del corpus, in cui i locutori possono parlare liberamente, e nella maggioranza dei casi si conoscono. A differenza di quanto avviene nelle conversazioni del CLIPS, però, nelle conversazioni faccia a faccia (LIP sez. A) i parlanti possono vedersi.

Nel suo complesso, il corpus LIP contiene approssimativamente mezzo milione di parole, cioè più o meno 25.000 parole per ogni tipo di testo e ogni città.

LIP	Testi A (minuti)	Testi B (minuti)	Ore
Firenze	175,41	141,21	5,27
Milano	205,68	188,64	6,55

Tabella 2: Dimensioni del campione del LIP esaminato. Il dato sul minutaggio è ricavato dal sito BaDIP (http://languageserver.uni-graz.at/badip/).

Per il confronto presentato in queste pagine sono state quindi esaminate 50.000 parole per Firenze e altrettante per Milano, un campione che, da questo punto di vista, è di dimensioni notevolmente superiori rispetto a quello tratto da CLIPS. Altri dati sulle dimensioni del campione tratto dal LIP sono raccolti nella tabella 2.

3.3 CLIPS e LIP: alcune differenze

Vale la pena di menzionare qui alcune differenze tra le conversazioni di CLIPS e del LIP, dato che, come si vedrà (cfr. §5 n.6), non si può escludere che esse abbiano conseguenze sulle scelte linguistiche dei parlanti, ed in particolar modo sul livello di formalità delle conversazioni.

In primo luogo, a differenza dei parlanti del LIP, i parlanti del corpus CLIPS sanno di essere registrati e non parlano di argomenti scelti liberamente, ma svolgono un gioco su richiesta di quanti stavano raccogliendo il corpus. Ciò non rende il parlato del CLIPS meno spontaneo di quello del LIP per quanto riguarda la pianificazione dello scambio, ma lo rende senz'altro più prevedibile nei contenuti e nel lessico.

Inoltre, nel caso dei *map task*, un tipo di conversazione orientata utilizzata in CLIPS, in cui ai parlanti è richiesto di ricostruire il percorso che uno dei due ha disegnato su un foglio, hanno ruoli pragmatici prestabiliti, poiché svolgono funzioni diverse nell'ambito del gioco (uno guida il gioco e l'altro viene guidato).

Infine, mentre CLIPS è un corpus molto controllato per quanto riguarda la provenienza geografica dei parlanti, poiché tutti coloro che sono registrati sono nati e cresciuti nelle città scelte come punti di

raccolta, nel LIP il controllo sulla provenienza regionale dei parlanti non sembra essere altrettanto stretto. Il LIP fotografa fedelmente la situazione linguistica italiana attraverso la registrazione di scambi fatta nelle diverse città scelte per la raccolta. Il fatto però che una conversazione abbia luogo a Milano, non implica che essa avvenga tra persone originarie della città, come mostra chiaramente questo esempio tratto da una conversazione telefonica della sezione milanese:

(7) 'U mittissi dentro 'u muturino? (LIP, Milano)

È chiaro che l'autore di questa battuta non parla una varietà settentrionale. I rari casi di questo tipo, in cui la provenienza del parlante è apparsa chiaramente eterogenea rispetto al resto corpus, non sono stati conteggiati ai fini dello studio[2].

4 Dislocazioni a destra in CLIPS e nel LIP

Lo spoglio ha fatto registrare la presenza di 286 DD nel corpus CLIPS dialogico e di 142 DD nel campione del LIP milanese e fiorentino. Nel complesso, quindi, sono state raccolte 428 occorrenze della costruzione in esame.
Nello spoglio sono state considerate congiuntamente sia le DD propriamente dette, realizzate in un'unità tonale eventualmente scandita al suo interno in due parti (Rossi, 1999), sia quelle in cui l'elemento topicale dislocato si presenta nettamente separato, come un ripensamento (Berruto, 1986) o *afterthought* (Lambrecht, 1981). In questi enunciati l'elemento coreferente con il clitico è aggiunto dal parlante dopo il completamento della frase, come una glossa

[2] Un'altra differenza tra il LIP e il CLIPS, interessante ma non direttamente rilevante per questo studio, riguarda la fluenza, espressa in termini di numero di parole grafiche prodotte per minuto nelle conversazioni (De Mauro *et al*, 1993: 44). Se si confrontano ad esempio la fluenza nei campioni fiorentini e milanesi dei due corpora, si nota che i campioni tratti da CLIPS presentano un numero di parole per minuto (p.p.m.) assai inferiore rispetto ai dialoghi del LIP (CLIPS: Milano, 66,69 p.p.m.; Firenze, 63,62 p.p.m.; LIP: Milano A, 126,24 p.p.m.; Milano B, 165,20 p.p.m.; Firenze A, 154,21 p.p.m.; Firenze B, 145,72 p.p.m.). La causa della minore "densità" dei dialoghi CLIPS rispetto a quelli del LIP potrebbe essere ravvisata nei tempi di riflessione necessari ai parlanti per risolvere i problemi posti dai compiti di ricostruzione del percorso e individuazione delle differenze.

esplicativa (cfr. es. (8)) o, appunto, come un'aggiunta con valore di ripensamento:

(8) Passi intorno al lupo lasciandolo dentro al percorso <pausa> il lupo (CLIPS-Firenze)

Gli enunciati con ripensamento occorrono piuttosto raramente nei campioni analizzati. Data l'esiguità del loro numero, sono stati conteggiati insieme agli altri nelle tabelle[3].

4.1 Tipi sintattici e modalità

Il tipo di DD di gran lunga più frequente nel campione di CLIPS è quella dell'oggetto diretto (circa il 90% delle occorrenze). Gli altri tipi di dislocazione sono assenti o sono presenti solo marginalmente.

Tipo	**Numerosità**	**Percentuale**
CLIPS		
cl + V + clausola	12	4,2 %
cl + V + partitivo	18	6,3 %
cl + V + oggetto indiretto	4	1,4 %
cl + V + oggetto diretto	251	87,8 %
Altro	1	0,3 %
Totale CLIPS	286	100 %
LIP		
cl + V + clausola	36	25,4 %
cl + V + oggetto indiretto	20	14,1 %
cl + V + oggetto diretto	86	61,6 %
Altro		
Totale LIP	142	100 %

Tabella 3: Tipi di DD in CLIPS e nel LIP
(Milano e Firenze, sezioni A e B): numerosità e percentuali.

[3] Nel corpus CLIPS la presenza degli *afterthought* è stata controllata esaminando le trascrizioni e ascoltando i dialoghi, ed è risultata essere sporadica (poco più di 10 casi). L'occorrenza di ripensamenti nel LIP è stata invece valutata solo sulle trascrizioni ortografiche. In questo campione, infatti, non è stato possibile controllare la presenza di ripensamenti attraverso l'ascolto dei dialoghi, poiché le registrazioni di questo corpus non sono disponibili. Per l'identificazione degli *afterthought*, comunque, dovrebbe essere sufficiente il solo trascritto: secondo (Berruto, 1986), infatti, l'*afterthought* è separato da una chiara pausa dal resto dell'enunciato. Ci si può aspettare, quindi, che tale marcata discontinuità prosodica, per la sua evidenza, sia indicata nella trascrizione con un segno di interpunzione o esplicitando altrimenti la presenza della pausa. Dall'esame del LIP, condotto nel modo appena descritto, i ripensamenti sembrano essere comunque molto pochi.

Anche nel LIP, sebbene in percentuale inferiore rispetto a CLIPS, la dislocazione dell'oggetto diretto è il tipo sintattico più rappresentato nel campione (oltre il 60% dei casi). Un altro quarto delle occorrenze è costituito da enunciati con ripresa clitica di una clausola subordinata (tabella 3).
Per quanto riguarda la modalità, in CLIPS la DD appare più frequentemente negli enunciati interrogativi che in quelli dichiarativi o iussivi. Il dato è rovesciato nel LIP, dove invece la maggior parte delle dislocazioni occorre negli enunciati dichiarativi (tabella 4).
Anche se con alcune difformità tra i due campioni esaminati, i dati confermano le tendenze generali osservate nella letteratura sulla DD: la costruzione ricorre spesso in enunciati interrogativi e il costituente ripreso è solitamente un oggetto diretto.

Modalità	**Numerosità**	**Percentuale**
CLIPS		
Dichiarativa	98	34,3 %
Interrogativa	180	62,9 %
Imperativa	7	2,4 %
Totale CLIPS	286	100 %
LIP		
Dichiarativa	93	65,5 %
Interrogativa	43	30,3 %
Imperativa	6	4,2 %
Totale LIP	142	100 %
Totale CLIPS / LIP	428	100 %

Tabella 4: Modalità delle DD in CLIPS e nel LIP
(Milano e Firenze, sezioni A e B): numerosità e percentuali.

4.2 Numerosità e distribuzione geografica: CLIPS

Esaminando più in dettaglio i dati sulla numerosità, si osserva che la distribuzione delle DD nel campione di CLIPS è diseguale dal punto di vista diatopico (tabella 5).
La costruzione è attestata in tutte le varietà campionate in CLIPS, anche se in misura diversa nelle varie città in cui è stato effettuato il rilevamento.

Città	Numerosità	Percentuale
Bari	13	4,6 %
Bergamo	6	2,1 %
Cagliari	25	8,8 %
Catanzaro	23	8,1 %
Firenze	31	10,9 %
Genova	25	8,8 %
Lecce	9	3,2 %
Milano	4	1,4 %
Napoli	13	4,6 %
Palermo	35	12,3 %
Parma	21	7,4 %
Perugia	30	10,5 %
Roma	23	8,1 %
Torino	12	4,2 %
Venezia	16	5,6 %
Totale	286	100 %

Tabella 5: Numerosità delle DD nella città del campione di CLIPS

Il dato più evidente riguarda la scarsità di attestazioni in area lombarda, a Milano e Bergamo, dove i casi di DD trovati nel corpus sono in totale 10. Il dato sulla Lombardia contrasta in modo abbastanza evidente con quello della maggior parte delle altre città. Nel nord-est (Venezia) e nel nord-ovest (includendo Genova oltre che Torino), la dislocazione appare abbastanza diffusa, ed essa è ben rappresentata anche in area centrale (Firenze, Perugia, Roma).
Nel meridione, e più in particolare a Napoli e in area pugliese settentrionale, la costruzione appare di meno, con percentuali un po' superiori al 4%; la percentuale scende ulteriormente in area salentina. La costruzione appare invece ben attestata in Sardegna e in area calabrese e siciliana. Palermo, con Firenze, presenta il numero di DD più alto del corpus.
Dall'esame del corpus CLIPS dialogico, quindi, emerge una differenza diatopica riguardante la diffusione della DD con ripresa clitica nelle diverse aree d'Italia: il dato di maggiore evidenza è la presenza poco più che sporadica della costruzione nei dialoghi dell'aerea lombarda.

Nell'interpretazione di questi dati, bisogna naturalmente tenere conto della possibilità che le differenze geografiche registrate, ed in particolare quella più evidente riguardante l'area lombarda, siano il risultato di caratteristiche idiosincratiche dei campioni esaminati. Infatti, benché il materiale raccolto in CLIPS per le diverse località sia facilmente confrontabile per numero di conversazioni e di parlanti (6 dialoghi e altrettanti parlanti per ciascun punto di rilevazione) e sia inoltre controllato anche dal punto di vista diafasico e diastratico, non si può escludere che il risultato riguardante la distribuzione geografica sia falsato da abitudini specifiche dei parlanti registrati, o da fatti occasionali dei quali non è possibile avere notizia. Di una certa importanza in questo contesto è il fatto che, nei campioni lombardi, le poche occorrenze della costruzione siano distribuite in modo diseguale tra i parlanti: nel caso di Bergamo tre parlanti su sei non utilizzano mai le DD, mentre a Milano la costruzione in esame è utilizzata solo da due parlanti. Nelle altre città, invece, la costruzione è, con pochissime eccezioni, attestata nelle produzioni di tutti i parlanti. Il fatto che in CLIPS sette parlanti lombardi su dodici non utilizzino mai la costruzione nei dialoghi rappresenta quindi un'interessante anomalia.
I dati di Bergamo e Milano contrastano in modo particolare con quelli di città come Palermo e Firenze. Per esaminare più approfonditamente il dato relativo all'aera lombarda, si è scelto di procedere ad un confronto mirato tra i dati fiorentini e milanesi. La scelta delle due località è dovuta soprattutto al fatto che Milano e Firenze sono tra le località campionate nel corpus LIP: scegliendole è stato quindi possibile ampliare di molto la base di dati.

4.3 I campioni milanesi e fiorentini del LIP e di CLIPS

Il confronto tra i dati ricavati dal corpus CLIPS e dal corpus LIP ha lo scopo di verificare se la differenza registrata nel CLIPS tra l'area lombarda e il resto d'Italia, esemplificate dalle città di Firenze e Milano, sia casuale o se sia invece significativa. A questo scopo sono stati confrontati i dati sulla numerosità ed è stato eseguito un test statistico per valutare la significatività delle differenze osservate.

4.3.1 Numerosità e significatività statistica

Per prima cosa sono stati confrontati i dati sulla numerosità (tabella 6). Confrontando i numeri assoluti e le percentuali di occorrenza della DD nei diversi sotto-corpora, la differenza di numerosità tra i campioni fiorentini e quelli milanesi è in generale confermata, anche se appare più evidente nel parlato *task-oriented* che nel parlato delle conversazioni non orientate.

	CLIPS		**LIP - Testi A**		**LIP - Testi B**	
Milano	4	11,4 %	34	43,6 %	28	43,8 %
Firenze	31	88,6 %	44	56,4 %	36	56,3 %
Totale	35	100 %	78	100 %	64	100 %

Tabella 6: Confronto CLIPS *vs* LIP: numerosità e percentuali

Per verificare che la differenza geografica osservata attraverso i diversi campioni non sia casuale è stato eseguito un test di Wilcox *one-sided*.
Prima di poter procedere con questo test è stato necessario normalizzare i dati del corpus CLIPS, in modo da poterli confrontare tra di loro e con i dati del LIP con maggiore accuratezza[4].
Come valore di riferimento per la normalizzazione è stata utilizzata la media delle occorrenze della DD nelle 15 località campionate in CLIPS. La distribuzione delle occorrenze all'interno di questo corpus è risultata essere normale, come mostra il risultato del test di Shapiro e Wilk (p-value = 0.729). La media delle occorrenze della dislocazione nel CLIPS (19.067 occorrenze per città) è stata quindi utilizzata come valore di riferimento per normalizzare i dati dei diversi campioni[5].
Per provare se il numero di occorrenze della DD a Milano sia significativamente inferiore rispetto al numero di occorrenze a Firenze attraverso tutti i campioni, è stato utilizzato il test di Wilcox *one-sided*, per il quale ci si è serviti dei valori di occorrenza della dislocazione nei vari campioni normalizzati con la media ricavata da

[4] Benché i campioni milanese e fiorentino di CLIPS siano largamente simili, sussiste comunque una differenza nelle dimensioni espresse in numero di parole grafiche (cfr. tabella 1: Firenze 5917 p.g., Milano 5002 p.g.).
[5] Si tratta infatti di un valore medio di occorrenze della dislocazione a destra ricavato da un numero molto ampio di varietà regionali: tale valore medio può essere quindi utilizzato come riferimento anche per l'esame di corpora diversi da CLIPS, soprattutto se in essi, come nel caso del LIP, sono rappresentate alcune delle varietà campionate in CLIPS.

CLIPS. Il risultato del test (p = 0.040) mostra che la frequenza della costruzione è significativamente inferiore a Milano rispetto a Firenze.
I dati tratti dai due corpora mostrano la stessa tendenza: i casi di DD con ripresa clitica, cioè, sono più numerosi nel parlato dei fiorentini che in quello dei milanesi. Dall'esame dei campioni di CLIPS e del LIP, quindi, è emersa una differenza diatopica significativa riguardante la diffusione di una costruzione solitamente attribuita, senza maggiori specificazioni, all'italiano comune. Indubbiamente la costruzione in questione è attestata in tutte le varietà geografiche esaminate, come mostrano i dati ricavati dal corpus CLIPS, ma essa non ricorre in tutte le varietà con pari frequenza: in particolare essa è significativamente meno diffusa in area milanese.

5 Dislocazione a destra e italiano standard

Per rendere conto di questa dissimmetria nella distribuzione diatopica della DD si può ipotizzare che le varietà milanese e fiorentina siano in rapporto parzialmente diverso con la varietà standard tradizionale normativa.
Nel 1985, Nora Galli de' Paratesi concludeva il suo saggio *Lingua toscana in bocca ambrosiana*, affermando che Milano era il centro dove l'italiano normativo era stato più ampiamente recepito. Milano è definito un centro non standardizzatore ma *standardizzato*, «in cui la lingua nazionale si è diffusa e affermata più solidamente» (p. 207). L'inchiesta prendeva in considerazione soprattutto fatti di pronuncia, ma suggerisce una pista che potrebbe aiutare nell'interpretazione di questi dati, e cioè che la maggiore standardizzazione dell'italiano milanese possa riguardare non solo la pronuncia, ma anche la minore accettazione di costrutti sintattici ritenuti tradizionalmente scorretti, come appunto le costruzioni "pleonastiche" come la DD con ripresa pronominale. La più profonda adesione alle norme dell'italiano tradizionale spiegherebbe quindi la minore frequenza della costruzione a Milano[6].

[6] I dati sulla numerosità esposti in §4.3.1 suggeriscono anche una possibile relazione tra livello diafasico e occorrenza della dislocazione a destra. Confrontando i dati percentuali nei tre campioni, infatti, la differenza tra Milano e Firenze sembrerebbe più marcata nelle conversazioni orientate (CLIPS) rispetto a quelle libere (LIP). Alcune delle differenze che distinguono i due corpora potrebbero avere conseguenze sul grado di formalità che i parlanti ritengono appropriato al contesto enunciativo. In

Secondo questa ipotesi, quindi, la differenza osservabile sul piano diatopico potrebbe essere spiegata guardando agli sviluppi storici recenti delle varietà regionali in questione.

6 Conclusioni

Il quadro sulla DD delineato in queste pagine si presta ad alcune considerazioni riguardanti la dinamica variazionale interna alla varietà media comune dell'italiano (l'italiano neo-standard/colloquiale o dell'uso medio).
Il tratto esaminato in queste pagine mostra, come si è visto, una distribuzione geografica parzialmente diseguale in diverse aree dell'Italia. L'analisi ha fatto registrare, quindi, una sorta di irregolarità nel tessuto della varietà comune, irregolarità che non riguarda la presenza/assenza di un dato tratto, quanto la sua centralità in una particolare area. In altre parole, l'irregolarità non concerne l'*attestazione* del tratto (la DD destra è chiaramente attestata a Milano), ma la sua effettiva frequenza nelle diverse aree regionali. La verifica sulla DD presentata in queste pagine, quindi, indica che i tratti tradizionalmente attribuiti alla varietà neo-standard/colloquiale o dell'uso medio, pur essendo attestati, possono presentarsi con frequenza significativamente diversa in aree diverse.
La presenza di irregolarità interne e sovrapposizioni tra la varietà media comune e altre varietà del repertorio italiano testimonia la

---

particolare, va ricordato che i dialoghi del LIP avvengono in contesti naturali, che i parlanti sono liberi di scegliere l'argomento di conversazione e che le registrazioni sono effettuate all'insaputa dei parlatori, a differenza di quanto avviene per CLIPS, dove i parlanti indossano cuffie e microfono per svolgere un gioco in un luogo ad essi estraneo. Aggiungo anche una considerazione sulla diastratia dei parlanti nei due corpora, di qualche importanza in questo contesto: mentre i parlanti registrati nel LIP non sono tutti uguali per età e grado di istruzione, quelli del CLIPS rappresentano un gruppo omogeneo costituito da giovani tra i venti e i trent'anni con istruzione universitaria. Quindi, l'informalità dei dialoghi del LIP, soprattutto di tipo A, in combinazione con un diverso rapporto con l'italiano normativo, potrebbe essere la ragione per cui la dislocazione emerge meno spesso nelle produzioni dei parlanti milanesi, soprattutto in contesti un po' più formali. I parlanti milanesi, cioè, controllerebbero di più le loro produzioni nel caso dei dialoghi CLIPS, poiché questi avvengono in un contesto più formale e sicuramente più artificiale, mentre lascerebbero emergere con più frequenza un tratto tradizionalmente ritenuto da evitare nelle conversazioni più informali del LIP.

complessità interna della varietà media. Nel caso della relativa non standard, cui si è fatto cenno nell'introduzione, la dinamica variazionale si presenta come tensione tra la varietà media, in cui tali clausole sono certamente attestate, ma non sembrano essere largamente diffuse, e le varietà diastraticamente (e diafasicamente) più basse. Nel caso della DD, invece, l'irregolarità è limitata ad un'area circoscritta e potrebbe essere dovuta al una tensione tra varietà media e lo standard tradizionale.

Riferimenti bibliografici

Albano Leoni, F. (2003) Tre progetti per l'italiano parlato, in N. Maraschio and T. Poggi Salani (eds) *Italia linguistica anno Mille. Italia linguistica anno Duemila*, Roma: Bulzoni, pp. 675-83.

Albano Leoni, F. and R. Giordano (eds) (2005) *Italiano parlato. Analisi di un dialogo* (Quaderni di Comunicazione parlata). Napoli: Liguori Editore.

Alfonzetti, G. (2002) *La relativa non-standard. Italiano popolare o italiano parlato?*. Palermo: Centro studi filologici e linguistici siciliani.

Anderson, A.H., M. Bader, E. Boyle, G. Doherty, S.C. Garrod, S. Isard, J. Kowtko, J. McAllister, J. Miller, C. Sotillo, H.S. Thompson and R. Weinert (1991) 'The HCRC Map Task Corpus'. *Language and Speech* 34, pp. 351-66.

Aureli, M. (2003) 'Pressione dell'uso sulla norma. Le relative non standard nei giudizi degli utenti'. *Studi italiani di linguistica teorica e applicata (SILTA)* XXXII, 1, pp. 45-67.

– (2004) Le relative non-standard in alcuni corpora di italiano parlato (LIP, LIR, LABLITA, AVIP), in F. Albano Leoni, F. Cutugno, M. Pettorino and R. Savy (eds) *Atti del Convegno nazionale "Il parlato italiano"* (Napoli, 13-15 febbraio 2003), Napoli: D'Auria Editore, CD-Rom: B02.pdf.

Aureli, M., M. Cennamo, C. Crocco and F. M. Dovetto (2005) Aspetti morfosintattici, in F. Albano Leoni and R. Giordano (eds) *Italiano parlato. Analisi di un dialogo*, Napoli: Liguori Editore, pp. 119-35.

Benincà, P. (2001²) L'ordine degli elementi nella frase e le costruzioni marcate. Ordine normale e ordini marcati, in L. Renzi, G. Salvi and

A. Cardinaletti (eds) *Grande grammatica italiana di consultazione* (I ed. 1988, Vol. I), Bologna: il Mulino, pp. 129-67.
Bernini, G. (1991) Frasi relative nel parlato colloquiale, in C. Lavinio and A.A. Sobrero (eds) *La lingua degli studenti universitari*, Firenze: La Nuova Italia, pp. 165-87.
Berruto, G. (1986) Le dislocazioni a destra in italiano, in H. Stammerjohann (ed.) *Tema-Rema in Italiano/Theme-Rheme in Italian/Thema-Rhema im Italienischen*, Tübingen: Gunter Narr, pp. 55-69.
– (1987) *Sociolinguistica dell'italiano contemporaneo*. Roma: La Nuova Italia Scientifica.
Bonvino, E. (2004) Fenomeni sintattici e prosodici relativi alla distinzione fra due tipi di soggetto postverbale, in F. Albano Leoni, F. Cutugno, M. Pettorino and R. Savy (eds) *Atti del Convegno nazionale "Il parlato italiano"* (Napoli, 13-15 febbraio 2003), Napoli: D'Auria Editore, CD-rom: F02.pdf.
– (2005) *Le sujet postverbal: Une étude sur l'italien parlé*. Paris: Ophrys.
Bozzone Costa, R. (1991) Tratti substandard nell'italiano di giovani adulti, in C. Lavinio and A.A. Sobrero (eds) *La lingua degli studenti universitari*, Firenze: La Nuova Italia, pp. 123-63.
Calabrese, A. (2001²) I pronomi personali (2. I pronomi clitici), in L. Renzi, G. Salvi and A. Cardinaletti (eds) *Grande grammatica italiana di consultazione* (I ed. 1988, Vol. I), Bologna: il Mulino, pp. 563-606.
Cardinaletti, A. (2002) 'Against optional and null clitics. Right dislocation vs. marginalization'. *Studia Linguistica* 56, 1, pp. 29-57.
Cordin, P. (2001²) Il clitico "ne", in L. Renzi, G. Salvi and A. Cardinaletti (eds) *Grande grammatica italiana di consultazione* (I ed. 1988, Vol. I), Bologna: il Mulino, pp. 647-55.
Cortelazzo, M.A. (2001) 'L'italiano e le sue varietà: una situazione in movimento'. *Lingua e Stile* XXXVI, 3, pp. 417-30.
D'Achille, P. (1990) *Sintassi del parlato e tradizione scritta della lingua italiana*. Roma: Bonacci.
De Mauro, T. (ed.) (1994) *Come parlano gli italiani*. Firenze: La Nuova Italia.
De Mauro, T. and M. Voghera (1996) Scala mobile. Un punto di vista sui lessemi complessi, in P. Benincà, G. Cinque, T. De Mauro and

N. Vincent (eds) *Italiano e dialetti nel tempo. Saggi di grammatica per Giulio Lepschy*, Roma: Bulzoni, pp. 99-129.

De Mauro, T., F. Mancini, M. Vedovelli and M. Voghera (1993) *Lessico di frequenza dell'italiano parlato (LIP)*. Milano: ETAS libri.

Fiorentino, G. (1999) *Relativa debole. Sintassi, uso, storia in italiano*. Milano: Franco Angeli.

Della Corte, F., C. De Santis, C. Panzieri and R. Vetrugno (2004) QOL I: per un bilancio quantitativo e qualitativo dei fenomeni innovativi nell'italiano contemporaneo, con una nuova ipotesi di lavoro, in A. Cardinaletti and F. Frasnedi (eds) *Intorno all'italiano contemporaneo. Tra linguistica e didattica*, Milano: Franco Angeli, pp. 17-33.

Galli de' Paratesi, N. (1985) *Lingua toscana in bocca ambrosiana. Tendenze verso l'italiano standard: un'inchiesta sociolinguistica*. Bologna: il Mulino.

Lambrecht, K. (1981) *Topic, antitopic, and verb agreement in non-standard French* (Pragmatics and Beyond). Amsterdam: John Benjamins.

Messina, S. (2007) L'"italiano vero-simile": la mimesi dell'italiano parlato nella fiction televisiva, in A. Elia and A. Landi (eds) *La testualità. Testo Materia Forme*, Roma: Carocci, pp. 187-228.

Mioni, A. (1983) Italiano tendenziale: osservazioni su alcuni aspetti della standardizzazione, in P. Benincà (ed.) *Scritti linguistici in onore di Giovan Battista Pellegrini* (vol. I), Pisa: Pacini, pp. 495-517.

Radtke, E. (1992) Varietà dell'italiano, in A. Mioni and M. Cortelazzo (eds) *La linguistica italiana degli anni 1976-1986*, Roma: Bulzoni, pp. 59-74.

Rossi, F. (1999) 'Non lo sai che ora è? Alcune considerazioni sull'intonazione e sul valore pragmatico degli enunciati con dislocazione a destra'. *Studi di grammatica italiana* 18, pp. 144-93.

– (2005) Tratti pragmatici e prosodici della dislocazione a destra nel parlato spontaneo, in E. Burr (ed.) *VI Convegno Società internazionale di linguistica e filologia italiana (SILFI), "Tradizione e innovazione"* (Quaderni della Rassegna, n. 43; Duisburg, 28 giugno-2 luglio 2000), Firenze: Cesati.

Sabatini, F. (1985) L'"italiano dell'uso medio": una realtà tra le varietà linguistiche italiane, in G. Holtus and E. Radtke (eds)

Gesprochenes Italienisch in Geschichte und Gegenwart, Tübingen: Gunter Narr, pp. 154-84.
Simone, R. (1997) 'Une interprétation diachronique de la «dislocation à droite» dans les langues romanes'. *Langue française* 115, pp. 48-61.
Sobrero, A.A. and C. Lavinio (eds) (1991) *La lingua degli studenti universitari* (Quaderni del GISCEL), Firenze: La Nuova Italia.
Sobrero, A.A. and I. Tempesta (2007) 'Scelta informatori e località', *CLIPS-Documento di progetto*, http://www.clips.unina.it/.

Predicazione nominale negli usi dell'italiano parlato
L'esempio della classe delle forme lessicali di "sentimento"

Emilio D'Agostino
Università degli Studi di Salerno
emidago@libero.it

Abstract

The following essay moves from the question of the lexicon-grammar and syntactic study of linguistic uses of spoken Italian. We believe that these kinds of researches have got at least two relevant scientific aspects. The first one is the definition of structural features of spoken sentences. The second one represents a more general goal as it concerns the comparison between the spoken sentence structures and those ones of writing Italian. The computational linguistics' software Nooj has been used to underline the difference between the huge Italian linguistic competence and its concrete, limited execution.

0 Premessa

È nostra opinione che, nelle indagini lessicali sulle caratteristiche degli usi scritti e degli usi parlati di una lingua, un'analisi di tipo Lessico-Grammaticale raggiunga un livello di informatività maggiore di una pura considerazione per lemmi bruti, in quanto esplicita gli usi sintattico-semantici delle varie forme e contribuisce a delineare la struttura predicativa degli usi scritti e parlati di una lingua.
Dal nostro punto di vista, nel parlato come nello scritto, l'unità minima del Lessico di una lingua non è una singola forma, ma una frase, o meglio ancora, un gruppo coerente di frasi accomunate dalle relazioni esistenti tra un determinato operatore ed i suoi argomenti. Relazione, questa, che va a inserirsi in reti funzionali che altrove abbiamo chiamato "costellazioni parafrastiche"[1]. Ed è in tal senso che, anche in questa occasione, riteniamo utile procedere attraverso il riconoscimento delle costruzioni, delle forme di frase, possibili, di una loro tipologia associata a dei lemmi e della loro distribuzione lungo frequenze d'uso diverse e lungo tipi o generi discorsivi differenti.

[1] Si veda D'Agostino (1992).

In seconda battuta, va notato che l'indagine di tipo harrisiano si concentra sui vincoli operanti sulla combinatoria linguistica e sulle restrizioni che influenzano l'equi-probabilità di occorrenza delle parole l'una rispetto all'altra. In tal modo, si individuano tre tipi di vincoli che veicolano ognuno un tipo di significato, in modo tale che si possa ritenere che il significato di una frase sia determinato direttamente dalle parole e dai vincoli: *un vincolo di ordine parziale*, *una distribuzione di probabilità non uniforme* e *una riduzione di forme fonemiche*. In particolare, *il primo,* quello di *ordine parziale*, organizza la struttura frasale, agendo tra occorrenze di parole negli enunciati. Si tratta di un ordine parziale sulle parole consistente in un ordinamento in cui alcune parole sono più in alto o più in basso rispetto ad altre, mentre altre non sono né più in alto né più in basso tra loro. In pratica, l'ordine parziale costituisce un vincolo sulle combinazioni delle parole: esso dice che nella posizione argomento prossima ad un dato operatore la probabilità di certe parole – quelle che non appartengono alla classe di argomenti per quell'operatore – è zero. In sintesi, ci si trova di fronte a determinazioni di tipo probabilistico.

D'altronde, una considerazione degli usi si mostra pertinente ogni qual volta si prende la strada dell'analisi del vocabolario di una lingua. Richiamando la distinzione tra lista delle "forme lessicali" e lista degli "usi lessicali" e sapendo che quest'ultima è significativamente più grande della prima, ci pare del tutto evidente come, intrapresa la strada dell'analisi lessicale dell'uso linguistico, si giunga sempre poi alla necessità di indagare nella direzione della selezione e della co-occorrenza. Essa è il risultato dell'applicazione sistematica dell'analisi sviluppata in termini di vincoli operanti sulle *regole di selezione e co-occorrenza* registrate per una lingua. In particolare, ci si riferisce ai citati "vincoli dell'ordine parziale" e "vincoli di verosimiglianza" operanti prima del processo di linearizzazione delle frasi, così come descritti in più occasioni da Z.S. Harris nella sua teoria formale della sintassi.

Già in due precedenti occasioni, in relazione all'analisi di due forme verbali dell'italiano di altissima frequenza, abbiamo avuto modo di sostenere che l'esplicitazione degli usi sintattico-semantici delle forme lessicali contenute in un corpus come quello del *Lessico di Frequenza*

dell'Italiano Parlato (LIP) consentiva di indagare meglio le variazioni possibili all'interno di un sistema linguistico[2].

1 Il Lessico Grammatica della Lingua Italiana

Le indagini Lessico-grammaticali hanno cercato di confrontarsi con corpora testuali significativi e ciò per due ragioni: *la prima*, di carattere applicativo, nella direzione dell'elaborazione di procedure di trattamento automatico di dati linguistici, *la seconda*, di carattere più generale, nella direzione della definizione della struttura predicativa di registri e usi diversi di una stessa lingua. Il programma di ricerca "Lessico-Grammatica della Lingua Italiana" (d'ora in avanti LGLI), ispirato da Maurice Gross a partire dal suo *Méthodes en syntaxe* del 1975, in circa venti anni di contributi, ha visto articolarsi il proprio sviluppo lungo le seguenti direttrici di analisi e descrizione tassonomica: frase semplice e operatori verbali, costruzioni a verbo supporto e operatori nominali, frasi idiomatiche e operatori composti, forme composte, usi del parlato, applicazioni della teoria degli automi a stati finiti e vocabolari di frequenza, di base e specialistici.

2 L'analisi della classe dei nomi di "sentimento"

Il *primo obiettivo* è quello di apportare dati empirici all'affermazione teorica della relativa predominanza della predicazione nominale negli usi del parlato dell'italiano contemporaneo, se confrontato con gli usi scritti. In ciò siamo confortati dalle conclusioni dei due nostri precedenti lavori sul LIP. In esse osservavamo come alcune linee di tendenza potessero rilevarsi e, allo stato attuale, ci pare che venga confermato il risultato al quale si era pervenuti nell'indagine su *fare* e *dare* qui citata. Anche in quest'occasione, infatti, appare come, almeno a questi livelli di frequenza, il ruolo delle forme nominali, in un'articolazione tipologica e strutturale complessa, sia più determinante di quanto il solo computo complessivo per lemmi facesse apparire. In particolare, le costruzioni a verbo supporto costituiscono una forma di modulazione sintattica privilegiata e, soprattutto, rappresentano una realtà statisticamente fortemente significativa e, quindi, il ruolo predicativo dei nomi ad esse correlate

[2] Si veda D'Agostino (1993, 2001).

appare esaltato. A parte l'uso di forme nominali che si realizzano esclusivamente in forme di frase a supporto (ad esempio: *bisogno*), nell'assoluta maggioranza dei casi analizzati, le costruzioni a supporto sono statisticamente dominanti rispetto agli altri usi concorrenti.
Il *secondo obiettivo* è quello di costruire una "grammatica locale" delle forme lessicali di "sentimento". In particolare, pensiamo alle procedure legate ai Dizionari Elettronici e a quelle fondate sugli Automi Finiti. Per quanto riguarda i primi, con Silberztein (1993) si riconosce che tre sono i caratteri fondamentali dei dizionari elettronici che li distinguono dai dizionari cartacei (oggi anche digitalizzati) comuni: (relativa) completezza, esplicitezza e legame con i programmi di trattamento automatico. La prima si giustifica sulla base della considerazione per la quale ogni perdita di informazione concernente le parole di un testo si trasforma in una potenziale perdita di capacità di analisi del testo. La seconda si giustifica sulla base del fatto che nulla è evidente e noto alla macchina, ragione per la quale tutto deve essere esplicitato e reso chiaro. Infine, tutte le informazioni accumulate devono essere in un formato coerente che segua sempre e comunque gli stessi standard, sia per quanto riguarda la forma dell'entrata lessicale, sia per quanto riguarda la sequenza delle informazioni ad essa associate. Per quanto concerne le applicazioni degli Automi Finiti, ricorderemo che un "automa finito" è un sistema che può assumere un numero finito di stati, uno di questi stati è distinto ed è definito "stato iniziale", un altro è definito "stato terminale". L'automa ha la proprietà di poter leggere, uno dopo l'altro, i simboli di una sequenza costruita su un certo vocabolario: quando l'automa ha letto un simbolo può passare allo stato successivo, cioè cambiare di stato, ma il nuovo stato assunto dipende sia dal simbolo letto sia dallo stato nel quale si trovava al momento della lettura. Gli automi finiti, quindi, consentono di rappresentare, in generale, dei linguaggi che sono costituiti da insiemi di sequenze riconosciute dall'automa, in particolare essi sono impiegati per rappresentare le grammatiche locali di soluzioni combinatorie rintracciabili nei testi trattati automaticamente. Per i richiami bibliografici rimandiamo a Gross (1997) e a Vietri (2008).
Il *terzo obiettivo* di questo contributo è quello di affrontare la questione relativa alla dissimmetria tra piano della forma e piano del senso in termini di un'indagine empirica fondata su un corpus del

parlato dell'italiano di dimensioni quantitative e qualitative importanti (LIP).
Infine: il campione è dato da forme lessicali di "sentimento". In particolare, si analizzano i lemmi *ansia* e *angoscia* e le forme a esse correlate (*ansia*, *ansioso*, *angoscia*, *angosciare*, *angosciarsi*, *angosciante*, *angoscioso*).

2.1 L'operatore *angoscia*[3]

È chiara la difficoltà di definizione semantica del campo lessicale in questione. Noi seguiremo la definizione di tipo distribuzionale fornita da Gross (1995) per il quale la classe è definita dalle seguenti caratteristiche di base: in prima battuta, presenza obbligatoria di un Soggetto Umano obbligatorio (*Numobbl*), stabilmente associato a un operatore nominale, possibilità di forme di frase a *Vsup* come *N_0umobbl avere N_1operatore*, *N_0umobbl essere Prep N_1operatore*, *N_0umobbl essere Nagg*, oppure *N_0operatore esserci in N_1umobbl*. In seconda battuta, la possibilità di una costruzione ad operatore verbale del *si V*. La micro-classe lessicale costituita da [*angoscia angosciare angosciarsi angosciato angoscioso angosciante angosciosamente*] è un buon esempio di quanto qui si sta sostenendo. Infine, possibilità di

[3] *Angoscia*, dal latino *angustia*(*m*) "strettezza" da *angere* "stringere"; la forma *angosciare* "respirare a fatica, con affanno" è attestata in I. da Lentini (*S'eo guardo, quando passo, inver' voi no mi giro, bella, per risguardare; andando, ad ogni passo getto uno gran sospiro ca facemi angosciare e certo bene ancoscio, c'a pena mi conosco, tanto bella mi pare*); allo stesso modo in Cino da Pistoia (*Omè amor, perché nel primo passo non m'assalisti sì ch'io fossi morto. Perché non dipartisti da me, lasso, lo spirito angoscioso ch'io porto?*). In quest'uso entra in italiano due-trecentesco nel campo delle conseguenze delle "pene d'amore" e così in Dante delle *Rime* (*ahi angosciosa e dispietata lima che sordamente la mia vita scemi*) e della *Vita nuova* (*questi pensieri, e li sospir ch'io gitto diventan ne lo cor sì angosciosi, ch'Amor vi tramortisce*). L'angoscia, che nella classificazione stoica delle passioni rientra nel campo dei "dolori" e in San Tommaso è la *tristitia*, con Dante della *Commedia* assume, oltre al senso di oppressione fisica, quello di oppressione dell'anima dovuta alla condizione di condanna divina per il male compiuto (Inf. XX: *Io era già disposto tutto quanto a riguardar ne lo scoperto fondo, che si bagnava d'angoscioso pianto*). In italiano contemporaneo, vari lemmi diversi sono associati: *angoscia*, *ansia* e *angustia*. In inglese si registra *anguish* (medio inglese *angwisshe*, attraverso il francese *angoisse*), *distress* e *dread*; in tedesco *Angst* e derivati; in francese *angoisse*, *angoisser*, *s'angoisser*, *angoissè* e il correlato *ansieux*; in spagnolo *angustiar* e derivati.

costruzione a completiva. In terza battuta, possibilità di dare luogo ad una costruzione di tipo causativo. In particolare, con quest'ultima caratteristica si può osservare l'eventuale dissimmetria tra piano del senso e piano della forma (es. *angoscia ~ ansia*).
Se si utilizza il principio della cosiddetta "espansione massima" (i.e. un operatore con tutti i propri argomenti), si può rendere conto delle relazioni trasformazionali che collegano, in tutta evidenza, le varie forme di frase e si può individuare l'espansione massima dell'operatore di "primo ordine". In particolare, si possono collegare le frasi con operatore nominale e aggettivale, queste con quella con operatore verbale di tipo pronominale intrinseco, e tutte con la costruzioni di tipo completivo. In tal senso, si ha un percorso di classi di equivalenza frastiche come il seguente:

A) operatore elementare O(n) [*angoscia angosciato angosciarsi*]
- frase nominale/aggettivale:
Num_x ha (l')angoscia
Num_x è in angoscia
Num_x è angosciato
- frase verbale con Vintr:
Num_x si angoscia

(B) aggiunzione di un ulteriore argomento costituito dagli equivalenti distribuzionali (*Num + F*):
(Num_y + F_y) fa che Num_x abbia (l')angoscia
(Num_y + F_y) fa che Num_x sia in angoscia
(Num_y + F_y) fa che Num_x sia angosciato
(Num_y + F_y) fa che Num_x si angosci
(Num_y + F_y) fa avere angoscia a Num_x
(Num_y + F_y) fa (?essere + stare) Num_x in angoscia
(Num_y + F_y) fa (essere + stare) Num_x angosciato
(Num_y + F_y) fa angosciare Num_x

(C) inserimento dell'operatore causativo O(oo) del tipo *causa*/*causare*:
(Num_y + F_y) è la causa del fatto che Num_x abbia (l')angoscia
(Num_y + F_y) è la causa del fatto che Num_x sia in angoscia
(Num_y + F_y) è la causa del fatto che Num_x sia angosciato
(Num_y + F_y) è la causa del fatto che Num_x si angosci
(Num_y + F_y) causa il fatto che Num_x abbia (l')angoscia
(Num_y + F_y) causa il fatto che Num_x sia in angoscia
(Num_y + F_y) causa il fatto che Num_x sia angosciato

(Num_y + F_y) causa il fatto che Num_x si angosci

(D) costruzione a completiva:
(Num_y + F_y) angoscia Num_x
associata alle due costruzioni aggettivali:
(Num_y + F_y) è angoscioso per Num_x
(Num_y + F_y) è angosciante per Num_x.

L'argomento con interpretazione "causativa" può assumere anche la forma di un complemento preposizionale, quindi, collocato a destra della sequenza:

(E) Costruzioni con secondo complemento:
Num_x ha (l')angoscia (a causa di + per + ecc.) (Num_y + F_y)
Num_x è in angoscia (a causa di + per + ecc.) (Num_y + F_y)
Num_x è angosciato (a causa di + per + ecc.) (Num_y + F_y)
Num_x si angoscia (a causa di + per + ecc.) (Num_y + F_y).

Le rappresentazioni in termini di automi a stati finiti dell'insieme di frasi sono le seguenti:

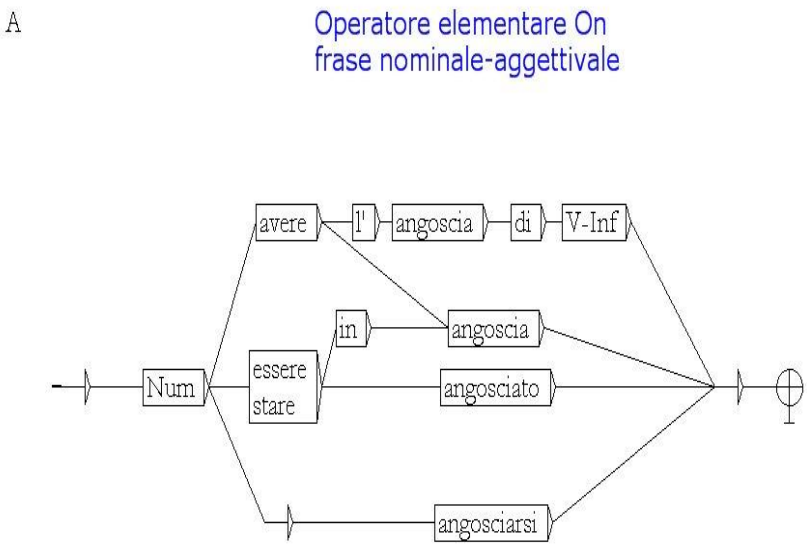


Figura 1

B

Aggiunzione di un ulteriore argomento
costituito dagli equivalenti distribuzionali (Num + F)

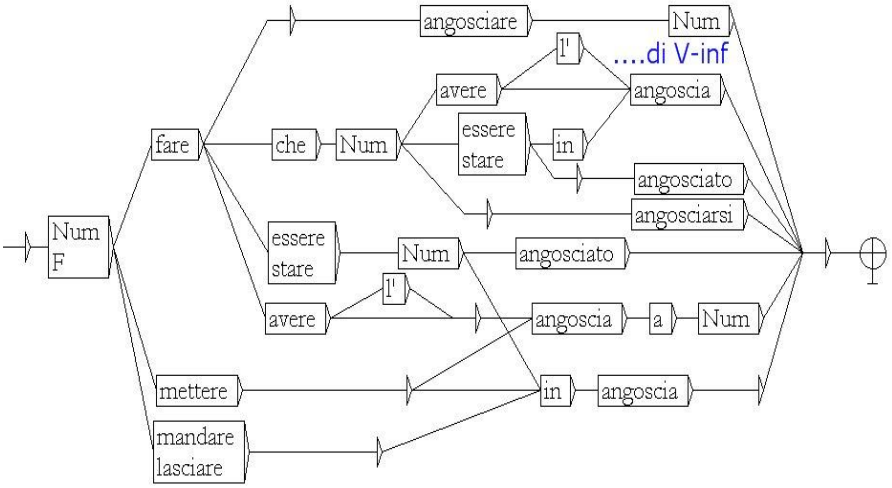


Figura 2

C

Inserimento dell'operatore causativo Ooo
del tipo causa/causare

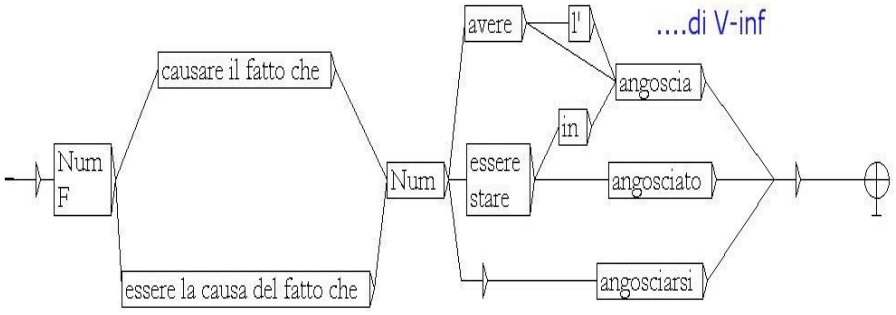


Figura 3

E

Figura 4

D

Figura 5

Con il riconoscimento dell'operatore nominale come operatore di primo ordine si ottiene non soltanto una descrizione più coerente, ma anche un risultato sulle tassonomie con il riequilibrio tra operatori

verbali, nominali e aggettivali, mettendo meglio in luce le relazioni fra essi.

2.2 Equivalenti distribuzionali

Il trattamento in termini di operatori nominali comporta automaticamente una maggiore attenzione ai *Vsup*. In particolare, si evidenziano le relazione di equivalenza stabilite dai supporti *avere*, *essere in* ed *essere*: *avere angoscia* ≡ *essere in angoscia* ≡ *essere angosciato* e delle loro varianti. Allo stesso modo si individuano gli equivalenti dell'operatore causativo verbale e nominale:

A) *avere* (*patire* + *soffrire* + *tenere* + *sentire* + *provare* + *ecc*.) *angoscia*
B) *essere in* (*patire di* + *soffrire di* + *stare in* + *ecc*.) *angoscia*
C*) essere* ((*stare* + *comportarsi in modo* + *avere un comportamento* + *mostrarsi* + ecc.) (*angosciato* + *angoscioso* + *angosciante*)).

La rappresentazione in termini di automa è la seguente:

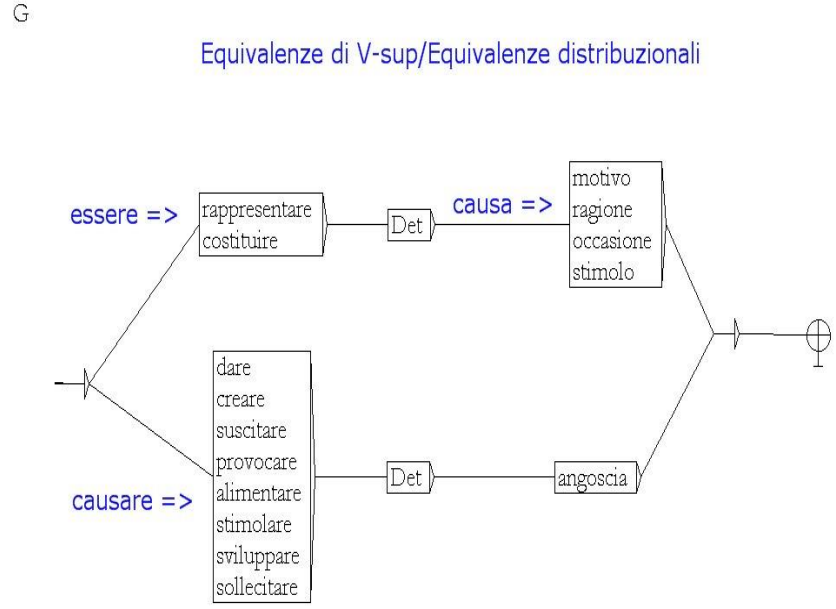


Figura 6

Inoltre, prendendo in considerazione la relazione di "conversione" (i.e. la realizzazione linguistica di una condizione logica)

dare/*ricevere*, si registra ancora una volta la relazione tipica delle costruzioni con tali operatori verbali, ad esempio nel caso degli operatori unari Onnn (costruzioni "dative"):

Sofia dà un libro a Eva ≡ Eva riceve un libro da Sofia.

Registrando anche l'equivalenza tra *ricevere*/*avere*:

N_y *ha angoscia da* N_x ≡ E*va ha angoscia da Max.*

2.3 I soggetti causativi

I soggetti delle costruzioni causative hanno sempre una duplice interpretazione che contrappone il tratto "volontario" a quello "involontario", come sotto-interpretazioni dell'opposizione "attivo-non attivo". Nel caso specifico della micro-classe qui analizzata, il carattere non-volontario e non-attivo dell'argomento aggiunto con l'inserimento dell'operatore di secondo ordine *causare* è legata, per definizione, dalla natura stessa dell'angoscia, sentimento del nulla, nella definizione di Sigmund Freud caratterizzata dal fatto di possedere «un carattere di indeterminatezza e di mancanza di oggetto». In tale direzione, va l'affermazione di Martin Heidegger:

> Col termine angoscia (*Angst*) non intendiamo quell'ansietà (*Ängstlichkeit*) assai frequente che, in fondo, fa parte di quel senso di paura (*Furcht*) che insorge fin troppo facilmente. L'angoscia è fondamentalmente diversa dalla paura. Noi abbiamo paura sempre di questo o di quell'ente determinato, che in questo o in quel determinato riguardo ci minaccia: la paura di… è sempre paura per qualcosa di determinato. E poiché è propria della paura la limitatezza del suo oggetto e del suo motivo, chi ha paura ed è pauroso è prigioniero di ciò in cui si trova. Nel tendere a salvarsi da questo qualcosa di determinato, egli diventa insicuro nei confronti di ogni altra cosa, cioè, nell'insieme "perde la testa". L'angoscia non fa più insorgere un simile perturbamento. È attraversata piuttosto da una quiete singolare. Certo, l'angoscia è sempre angoscia di…, è sempre angoscia per…, ma non è per questo o per quello. Tuttavia, l'indeterminatezza di ciò di cui e per cui noi ci angosciamo non è un mero difetto di indeterminatezza, bensì l'essenziale impossibilità della determinatezza (Heidegger, 1929: 67).

2.4 Ansia e angoscia

L'appartenenza alla stessa classe di senso non implica l'appartenenza contestuale alla stessa classe di comportamento sintattico. Un esempio semplice è dato dai lemmi seguenti nel loro uso di operatori verbali elementari applicati su tre argomenti: *fornire*, *dare*, *privare* e *togliere*. Si osserverà che, pur essendo le stesse le relazioni tra gli argomenti per *fornire* e *dare* da un lato e, dall'altro, per *privare* e *togliere*, dall'altro, i comportamenti sintattici sono differenti: *fornire q.cn di q.cs*, *dare q.cs a q.cn*, *privare q.cn di q.cs* e *togliere q.cs a q.cn*. Con analoga dissimmetria si comportano usi verbali di tipo "verbo supporto" (*Vsup*) come *dare*, *ricevere* e *avere*: *dare uno schiaffo a q.cn*, *ricevere uno schiaffo da q.cn* e *avere uno schiaffo da q.cn*.
I due lemmi *angoscia* e *ansia*, che nel Dizionario di Tullio De Mauro sono definiti in modo pressoché sinonimico:

angoscia
1a. AU stato di agitazione, affanno, sofferenza psicologica che può generare anche manifestazioni patologiche: *sentire*, *avere una grande angoscia*; *essere in angoscia*
1b. OB difficoltà di respiro
2. TS psic. ansia
3. TS filos. angoscia esistenziale
ansia
1a. stato di agitazione dovuto a timore, incertezza o attesa di qcs.: *stare in ansia per qcn.*
1b. desiderio, brama: *ansia di gloria*
2. TS psic. reazione emotiva, causata da conflitti inconsci, sensazioni di incertezza, timore, oppressione, ecc. di fronte a qcn. o qcs. percepito come pericolo

in realtà presentano comportamenti sintattici differenti in gran parte. Infatti, se entrambi possono reggere un'infinitiva con soggetto coreferente con il soggetto della frase con operatore nominale:

A) *Sofia ha l'ansia di partire*
B) *Sofia ha l'angoscia di partire*

anche se l'interpretazione di (A) è duplice[4], soltanto il secondo, come s'è visto, è correlato ad un operatore verbale intransitivo del tipo *si V*:

C) *Sofia* ((**si ansia* + *si angoscia*) (*per il partire* + *per la partenza*))

e all'operatore causativo di tipo verbale:

D) *partire* (**ansia* + *angoscia*) *Sofia*

3 *Angoscia* e *ansia* nel LIP

Nel corpus del LIP sono registrate le seguenti occorrenze dei due lemmi:

1	2	3	4	5	6
M	**A**	**11**	**12**	**A**	ecco che spengono anche la luce ahi * ho paura # ragazzi che ansia % %

1	2	3	4	5	6
N	**C**	**9**	**86**	**B**	* le poesie di guerra quelle che noi abbiamo studiato sono Veglia e San Martino del Carso in queste due poesie si vede tutta l'angoscia del diciamo del del poeta
R	**A**	**1**	**192**	**A**	sì eh ma questo è un po' triste a me mette un'angoscia * eh vabbe' insomma
R	**A**	**4**	**53**	**D**	basta $ poi non lo so uno s'angoscia tutto il giorno con la chimica la sera vorrebbe pure svagarsi trastullarsi con altre cose
M	**B**	**46**	**115**	**A**	ma non fare quelle superpossessive che sono diventate angoscianti
M	**C**	**6**	**255**	**A**	un po' angosciato certo perché appunto sa
M	**C**	**6**	**22**	**F**	cosa che lui in quel suo momento di gioia angosciosa angosciosa faceva sempre
M	**C**	**6**	**23**	**G**	cosa vuol dire gioia angosciosa *
M	**C**	**6**	**24**	**A**	vuol dire che era felice gioia perché era felice perché sapeva di stare salendo in camera e poteva dormire con i bambini angosciosa però perché sapeva anche che per riuscire a salire doveva

[4] Ciò si rivela grazie all'applicazione dell'operatore aggettivale *ansioso*: *Sofia è ansiosa di partire* = "Sofia non vede l'ora di partire" differente da *Sofia soffre di ansia per* (*il partire* + *la partenza*).

Se queste sono le occorrenze dei lemmi *ansia* e *angoscia*, vanno notate alcune loro caratteristiche. *In primo luogo*, è evidente come essi non appartengano ai ranghi alti della frequenza, in particolare il primo. *In secondo luogo*, va notato come, pur in bassa frequenza, il secondo vede realizzarsi la maggior parte delle sue possibilità: da operatore elementare nominale, a verbo pronominale, a costruzione causativa. *In terzo luogo*, si può osservare che in entrambi i casi, se si escludono gli enunciati che sono indubbiamente inseriti in conversazione tipo "faccia a faccia con presa di parola non libera" (es. "esami") *ansia* e *angoscia* sono contraddistinti da un uso generico o improprio, indicando un vago sentimento negativo. In realtà, i lemmi concernenti "emozioni" e "sentimenti", in italiano contemporaneo sembrano essere impiegati, nell'uso quotidiano, in tal modo "sfumato". Esempio ne sia il lemma *paranoia*, così come si registra nello stesso LIP:

1	2	3	4	5	6
F	B	14	381	A	* mannaggia questo di qui m'ha m'ha m'ha mandato in paranoia con queste telefonate

4 Anemia linguistica e *canniledi di picuraru*

Devo chiarire in partenza cosa qui intendo con l'espressione "usura semantica" e con il termine "svuotamento" associato all'uso che ho or ora detto "sfumato". In un caso come nell'altro, sulla scia di Tullio de Mauro, voglio indicare quel fenomeno realizzatosi nel processo di standardizzazione della lingua italiana durante lo sviluppo della società dei consumi grazie all'uso e all'abuso di *parole-esca*, *parole-precetto*, *slogan* che ha prodotto la rimozione del contenuto della classe di sensi – cioè del significato – associata a una parola o a un'espressione[5]. Credo che tale nozione sia ben distinta da quella di "uso generico" degli elementi lessicali di una lingua e che essa stia a indicare la progressiva discesa verso qualsiasi stagione culturale e linguistica: qualcosa di più grave di un semplice impoverimento o di un uso improprio. Spesso tale svuotamento ha significato soltanto assenza di regole e di principi etici generatori di quella confusione linguistica che si è prodotta almeno da un paio di decenni. Interessata

[5] Corti (2001) utilizza a tale riguardo l'espressione "anemia linguistica". Per la definizione di "significato" come "classe di sensi" rimando a Prieto (1967).

o casuale questa ha prodotto una "omogeneizzazione" culturale e politica[6]. Tale confusione, ben distinta da quella babelica, ha comportato la perfetta opacità di tutta una classe di parole divenute "usurate". Vuote perché i loro sensi sono divenuti ignoti nell'uso quotidiano: Leonardo Sciascia direbbe che i *canniledi di picuraru* sono ormai ignoti agli stessi pecorai[7].

5 Conclusioni

Lo studio lessicale e sintattico degli usi dell'italiano parlato possiede, a nostro avviso, almeno due valenze rilevanti. La prima di ordine specifico: la definizione delle proprie caratteristiche strutturali. La seconda più di ordine generale: il confronto con quelle proprie dell'uso scritto. Nel nostro caso, crediamo di aver contribuito a tale obiettivo attraverso l'impiego di strumenti automatici che evidenziano la differenza esistente tra le possibilità del sistema lingua italiana e le sue effettive realizzazioni.

Riferimenti bibliografici

Corti, M. (2001) *Nuovi metodi e fantasmi*. Milano: Feltrinelli.

D'Agostino, E. (1992) *Analisi del discorso. Metodi descrittivi dell'italiano d'uso*. Napoli: Loffredo.

– (1993) *Sociolinguistica computazionale. Un'applicazione descrittiva al corpus del LIP* Salerno: Ed. 10/17.

– (2001) *Le forme lessicali del parlare: analisi quantitativa e qualitativa del parlato italiano.* Napoli: Ed. Scientifica.

Freud, S. (1926) *Hemmung, Symptom und Angst.* Wien: Internationaler Psychoanalytischer Verlag [trad. it. (1988) *Inibizione, sintomo e angoscia*. Milano: Boringhieri].

[6] Al termine *omologazione* usato da Pier Paolo Pasolini preferisco *omogeneizzazione*, in quanto esso esprime meglio quanto oggi è registrabile: una baumaniana collosa pappa, in cui moltissimi discorsi possono significare tutto e il contrario di tutto, dal colore indefinito come quello, per l'appunto, degli omogeneizzati.

[7] Il riferimento alle "candeline del pecorario" è legato, per l'appunto, alle modificazioni rapidissime subite dalla lingua italiana più volte riprese da Pier Paolo Pasolini. Sciascia, anche se in relazione al caso-Moro, vi accenna in *L'affaire Moro*, 1994, Adelphi, Milano.

Gross, M. (1997) The Construction of Local Grammars, in E. Roche and Y. Schabès (eds) *Finite-State Language Processing*, Cambridge, MA: MIT Press, pp. 329-54.
Heidegger, M. (1929) *Was ist Metaphysic*, [trad. it. (1987) *Che cos'è la metafisica*?, Milano: Adelphi].
Prieto, L.J. (1967) *Principi di noologia*. Roma: Ubaldini [ed. or. (1964) *Principes de noologie*. The Hague: Mouton].
Silberztein, M. (1993) *Dictionnaires électroniques et analyse automatique de textes: le système INTEX*. Paris: Masson Ed.
Vietri, S. (2008) *Dizionari elettronici e grammatiche a stati finiti*. Salerno: Plectica.

Costrutti con *c'è* e nome proprio in una telecronaca sportiva: configurazioni funzionali e valori testuali[*]

Nunzio La Fauci[♠], *Heike Necker*[♠], *Sophia Simon*[♠], *Liana Tronci*[♦]
Universität Zürich[♠], Università per Stranieri di Siena[♦]
lafauci@rom.uzh.ch; hnecker@rom.uzh.ch; sosimon@rom.uzh.ch; tronci@unistrasi.it

Abstract

This paper is part of a larger research on certain morphosyntactic aspects of a peculiar kind of "spoken language" in Italian, i.e. live broadcasts of soccer matches. The report we take into account concerns the match Germany vs. Italy played in Dortmund, July the 4th 2006. In this report, which is fully transcribed, one finds a number of constructions such as C'è Perrotta che innesta la quarta, C'è Totti al centro, C'è Schneider, *namely syntactic structures with combinations between* essere *'to be' and* ci *'there', which involve proper names. How are these constructions characterized from a syntactic point of view? What is their textual relevance? Our paper aims at providing answers to both questions.*

1 Introduzione

La sera del 4 luglio 2006, a Dortmund, ebbe luogo una delle due partite semifinali del Campionato mondiale organizzato dalla Federazione calcistica tedesca. La squadra dei padroni di casa vi affrontò la Nazionale italiana. L'incontro ebbe una durata di poco più di due ore. Alla fine dei tempi regolamentari a nessuna delle due squadre era infatti riuscito di segnare un goal e dalla sfida doveva necessariamente venire fuori una delle due finaliste che si sarebbero contese il titolo mondiale a Berlino di lì a pochi giorni. Terminati i tempi supplementari, l'esito fu favorevole ai colori italiani: la Germania uscì sconfitta per 0 a 2.
All'importante evento sportivo, la prima rete della televisione pubblica italiana dedicò una telecronaca in presa diretta, curata da

[*] La ricerca di cui rende conto questo scritto è stata sviluppata congiuntamente dagli autori. Si precisa tuttavia per le esigenze legali che la stesura dei paragrafi 1 e 6 si deve a NLF e a LT quella di 2 e 4, mentre SS è responsabile del paragrafo 3 e HN del 5. Grazie ad Andrea Bonazzi e Ignazio Mirto per il loro aiuto.

Marco Civoli[1], con la collaborazione di Sandro Mazzola, come commentatore tecnico[2]. Per poco più di centoventi minuti, a supporto delle immagini, Civoli e Mazzola descrissero quanto avveniva sul campo per milioni di telespettatori italiani, informandoli inoltre su aspetti collaterali e discutendo delle fasi salienti del gioco.
Quanto fu detto da loro in quelle due ore è stato raccolto e trascritto dagli autori della presente nota per scopi di analisi sintattica: un testo orale dalla natura certamente particolare[3]. In specifico riferimento al tema su cui qui si pone l'accento, si è così potuto rilevare che nel testo a due voci della telecronaca ricorrono numerosi costrutti in cui una forma finita del verbo *essere* accompagnata dalla particella *ci* è riferita ad un nome proprio, che nel séguito sarà chiamato *pivot*: in media, quasi due ricorrenze ogni tre minuti. Ecco alcuni esempi rappresentativi[4]:

(1) L'intervento di Friedrich # la palla in gioco # *c'è Schneider # c'è Borowski però in mezzo al campo che è libero*
(2) [Gattuso] allarga sapientemente il gioco # *c'è Zambrotta # c'è Camoranesi sulla sua corsa e poi scivola*
(3) Scatta anche Toni # *c'è Totti al centro # c'è Perrotta che innesta la quarta*

In questi passaggi, un *c'è* incalza l'altro, come si vede, e non sono certo i soli nella telecronaca così fatti. Assistendovi e ascoltandone il testo con la distanza e l'intenzione di osservatori spassionati e non (più) emotivamente coinvolti dalle vicende di cui vi si riferisce, si ha anzi l'impressione di trovarsi di fronte allo stilema di

[1] Nel séguito, ove necessario, indicato come A, quando si citano passaggi del testo della telecronaca.
[2] Indicato come B, ove necessario. Citazioni dal testo della telecronaca prive di indicazioni sono *per default* da riferire ad A.
[3] V. in proposito La Fauci and Necker (2010) and Necker, Simon and Tronci (2009). Al primo dei due lavori si rimanda inoltre per una concisa riflessione sugli aspetti linguistici della telecronaca sportiva e per un breve rendiconto dei criteri di trascrizione del testo, che fa da base sperimentale anche della ricerca di cui rende conto questo scritto.
[4] In questi esempi e in quelli che seguiranno, le parti pertinenti per la specifica argomentazione sono poste in corsivo.

un'estemporanea (ma non perciò asistematica) composizione linguistica orale. Come stilema, esso cade peraltro perfettamente sotto le note riflessioni di Roman Jakobson sulle procedure iterative e sul loro ruolo nella costruzione della parola poetica, tanto orale quanto scritta[5].

La circostanza è naturalmente idiosincratica, nei suoi caratteri fenomenici macroscopici, ed ascrivibile per questa ragione alla *parole*. La *parole* non è però che una faccia dell'oggetto della ricerca linguistica: in essa e nei suoi accidenti il sistema fenomenicamente si realizza e ciò significa che in essa e tra i suoi accidenti il sistema si manifesta. Nei suoi caratteri idiosincratici, la telecronaca in questione è di conseguenza un contesto di studio privilegiato di costrutti con *c'è* e forme simili, di cui la letteratura sulla sintassi dell'orale italiano si è da tempo occupata.

Nella sua analisi, questa nota, breve e servile, procede da costrutti diversamente complessi verso il costrutto più semplice tra quelli rilevati nel suo àmbito osservativo, allo scopo di recare al dibattito scientifico sul tema un'elementare e desultoria panoramica dei dati a loro modo significativi che vi ha reperito[6]. Essa s'interroga sulle configurazioni funzionali che tali dati manifestano e sui valori sintattici e testuali che essi testimoniano. Non prende tuttavia un indirizzo tassonomico pregiudiziale che non sia il più banalmente grammaticale.

2 Ricorrenze con modificazione del *pivot*

Il modulo *c'è Perrotta che innesta la quarta* (es. 3) è quello su cui, com'è noto, si è soprattutto concentrata la letteratura sugli aspetti sintattici dell'italiano parlato[7]. In esso, il *pivot* funge da elemento

[5] In particolare Jakobson (1973 [1966]).

[6] Per ciascuno dei differenti moduli strutturali descritti nelle pagine che seguono si forniscono tutte le attestazioni presenti nella telecronaca.

[7] Per es., già Sabatini (1980: 80) osservò l'esistenza di una costruzione *c'è N che V* in cui «il verbo è accompagnato […] dalla particella *ci*, che semplicemente attualizza (e non propriamente localizza) il dato» e vi individuò una caratteristica morfosintattica dell'"italiano parlato unitario". Sulla sua scorta, e tra gli altri, Berruto (1986), Berretta (1994), Venier (2002), De Cesare (2006) hanno concentrato la loro attenzione sul fenomeno, ponendolo sotto l'etichetta di *c'è* presentativo. La

argomentale di una modificazione di forma variabile ma dalla funzione uniformemente attributiva. Nella telecronaca, le attestazioni di tale modulo, senza essere molto frequenti, sono tuttavia in numero significativo:

(4) [Pirlo] allarga sulla destra dove c'è Zambrotta # *c'è Camoranesi che prova lo scatto*
(5) *c'è Totti adesso che prova a far viaggiare Toni*
(6) *c'è Totti che rincorre questo pallone* # c'è anche Schneider # il *tackle* fra i due
(7) il tocco verso Camoranesi # *c'è Totti che attende*
(8) Camoranesi # bisogna allargare il gioco adesso # *c'è Grosso che ha capito* e *c'è Totti che lo vede*
(9) parte Perrotta # prova a servirlo # *c'è Mertesacker che ha compreso tutto*
(10) e *c'è Gattuso però che ha anticipato le mosse di Ballack*
(11) scatta anche Toni # c'è Totti al centro # *c'è Perrotta che innesta la quarta* (= es. 3)
(12) parte Del Piero # posizione regolare # *c'è Grosso attaccato alle spalle da Odonkor*

Ai casi semplici sono poi da aggiungere quelli sintatticamente più complessi, per la presenza, oltre alla modificazione attributiva, di ulteriori predicazioni manifeste, dalla natura funzionale non-attributiva, come *largo* in (13), *in mezzo al campo* in (14), di cui si dirà poco sotto:

(13) adesso *c'è largo Zambrotta che riceve*
(14) c'è Schneider # *c'è Borowski però in mezzo al campo che è libero* (= es. 1)

Allo stesso modulo, ma in maniera superficialmente meno esplicita,

---

designazione pare del resto una specializzazione di quella proposta da Salvi (1988: 127), per il quale più generalmente «dal punto di vista semantico *esserci* è la variante presentativa dell'uso locativo di *essere*». Nella letteratura, la si trova tuttavia riferita, in modo ancora una volta specializzato, per costrutti formalmente caratterizzati in modo diverso: Schiller (1995) l'adopera per es. per le costruzioni introdotte da *ecco*.

sono poi da riferire probabilmente anche i casi in cui il *pivot* è ripreso in ciò che segue immediatamente per via anaforica:

(15) c'è Zambrotta # *c'è Camoranesi sulla sua corsa e poi scivola* (= es. 2)
(16) si smarca Toni # *c'è Camoranesi # lo vede* # servito Toni
(17) arriva anche Totti # Metzelder # *c'è Zambrotta # prova lo stacco*
(18) ma *c'è Pirlo # con estrema calma esce dalla nostra area di rigore*
(19) parte adesso Grosso # lo vede # *c'è anche Pirlo # deve fermarsi*
(20) quindi Gattuso # *c'è Grosso # viene attaccato da Schneider*

3 Ricorrenze con predicazione locativa manifesta

Una predicazione che si presta ad un'interpretazione spaziale e rispetto alla quale il *pivot* funge da argomento localizzato pare la più banalmente combinabile con il modulo qui al centro dell'attenzione e, in effetti, numerose e banali ne sono le ricorrenze. In esse (come è comprensibile) la predicazione locativa precede tanto la forma verbale finita, quanto il *pivot*:

(21) quindi Pirlo Totti ancora per il regista del Milan # allarga sulla destra *dove c'è Zambrotta*
(22) Borowski, giocatore del Werder Brema # *vicino a lui c'è Philipp Lahm*
(23) il rilancio sempre per Klose # *su di lui c'è Materazzi*
(24) parte il traversone per Friedrich # *da quella parte c'è Fabio Grosso*
(25) di prima a cercare Camoranesi # *da quella parte c'è Lahm*
(26) rincorre il pallone poi Borowski # *su quella palla c'è anche Camoranesi*
(27) il rilancio di Lehmann # nella zona di Borowski # *alle sue spalle c'è Camoranesi*
(28) lo vede # è in posizione regolare # *alle spalle c'è Lahm*
(29) adesso Friedrich # *da quella parte c'è Materazzi* # in anticipo Perrotta

(30) dentro Camoranesi # *vicino a Lehmann c'è Luca Toni*
(31) Lahm *sulle cui piste c'è Zambrotta*
(32) *al centroarea c'è Perrotta*
(33) *dentro nello spazio c'è Camoranesi*
(34) Friedrich # in direzione di Podolski # *dall'altra parte c'è Schweinsteiger*
(35) lo affronta Gattuso cerca di mandarlo fuori giri # *da quella parte c'è anche Materazzi*
(36) in direzione di Gilardino # *dalle parti di Iaquinta c'è Lahm*

Sono in numero minore e sono, in modo coerente, meno banali dal punto di vista enunciativo-comunicativo quelle costruzioni in cui la specificazione predicativa segue forma verbale e *pivot* e in cui, di conseguenza, quest'ultimo si atteggia a topico:

(37) Toni però si è allargato troppo # *adesso c'è Camoranesi in mezzo* # parte il cross
(38) è bravo poi ad impedire la conclusione ancora di Schneider # *c'è Friedrich sul fronte opposto*
(39) largo anche Borowski # servito # *c'è Zambrotta da quella parte*
(40) la palla in gioco # c'è Schneider # *c'è Borowski però in mezzo al campo che è libero* (= ess. 1, 14)
(41) allarga sapientemente il gioco # c'è Zambrotta # *c'è Camoranesi sulla sua corsa* e poi scivola (= ess. 2, 15)
(42) scatta anche Toni # *c'è Totti al centro* # c'è Perrotta che innesta la quarta (= es. 3)
(43) Totti # *c'è Grosso sulla sinistra* # cerca il tocco verso Gilardino che mantiene il possesso di palla

4 Ricorrenze con complemento predicativo

Quanto a configurazione grammaticale, non vanno confusi né con gli uni né con gli altri, però, quei costrutti in cui il *pivot* funge sì da elemento argomentale di una predicazione dall'interpretazione locativa, ma in cui tale predicazione si presenta, come parte del discorso, come un aggettivo. Ci si riferisce ai casi non numerosi ma preziosi di:

(44) il tocco verso Totti # che arretra # *adesso c'è largo Zambrotta* che riceve ma attenzione perché c'è Grosso # lo ha visto (= es. 13)
(45) il cambio di passo di Zambrotta # *largo c'è Totti*
(46) Grosso si rifugia sul sinistro di Materazzi # *c'è Iaquinta larghissimo* # non lo contrasta nessuno

La possibilità dell'aggettivo di ricorrere tanto dopo il *pivot* (es. 46) quanto prima del *pivot* – e in quest'ultimo caso prima della forma verbale (es. 45) e, soprattutto, tra essa e il *pivot* (es. 44) – dice, da un lato, che tale aggettivo non è un semplice modificatore del *pivot*, dall'altro, che, se la sua funzione è, rispetto al *pivot*, di predicato e se quindi il *pivot* ne è argomento, la relazione predicativa in questione si realizza in una configurazione complessiva di unione proposizionale (cfr. Davies and Rosen 1988)[8].
L'evidenza è meno dirimente ma non è tuttavia da escludere che sia attribuibile al medesimo modulo grammaticale anche:

(47) palla recuperata da Materazzi # c'è Totti # ancora Camoranesi # frenato dal *tackle* perfetto di Kehl # mentre *c'è Gattuso a terra* per crampi

Lungi dall'essere un supplemento locativo del tipo esemplificato nel paragrafo precedente, in questo esempio la forma *a terra* si caratterizza come manifestazione di un predicato interno di una unione proposizionale: si tratta formalmente di aggettivo composto da interpretare come 'in una momentanea condizione fisica tale da non poter partecipare al gioco'. *A terra c'è Gattuso*, *c'è a terra Gattuso*, anche se non attestate nella telecronaca, paiono d'altra parte sequenze più che plausibili.
Altrettanto, se non più ragionevole è attribuire al tipo sintattico in questione la doppia ricorrenza (una volta col verbo all'imperfetto: è B

[8] In correnti e alternativi quadri teorici, ci si riferisce al fenomeno con l'etichetta di *small clause*: insieme con il *pivot*, i predicati *largo* e *larghissimo* sono appunto uno *small clause*, inglobato nella configurazione sintattica complessiva della proposizione.

che parla e sulla questione si tornerà; una volta col verbo al presente) che si rileva in:

(48) **A:** sì # [Klose] ha fatto 10 reti quest'anno nella Bundesliga # buon tiratore ma in questo caso da quaranta metri ha cercato veramente il jolly **B:** *poi in porta c'era Buffon* eh **A:** *in porta c'è Buffon* # perfetto nell'uscita su Klose

Lungi dall'essere una predicazione locativa, come pure superficialmente si presenta, *in porta* vale 'come portiere', 'nella funzione di portiere' e con tale funzione la variante *tra i pali* ricorre, riferita tanto al portiere della squadra nazionale tedesca quanto a quello dell'italiana, nelle battute iniziali della telecronaca, quando sono presentati i protagonisti dell'incontro:

(49) *Lehmann tra i pali* Friedrich Mertesacker Metzelder Lahm Kehl Schneider Borowski Ballack Klose e Podolski
(50) E adesso i nostri ragazzi # *Buffon tra i pali* Zambrotta Cannavaro Materazzi # intanto abbraccio e bacio tra i due capitani Cannavaro e Ballack # dicevamo Buffon Zambrotta Cannavaro Materazzi Grosso Pirlo Gattuso Perrotta Camoranesi Totti Toni

5 Il grado zero

Privo di ogni altro elemento rilevante per la sua sintassi interna, il modulo *c'è Schneider* (es. 1) o *c'è Zambrotta* (es. 51) è, tra quelli pertinenti per la presente ricerca, certamente il più semplice e ricorre con rilevante frequenza nel testo della telecronaca:

(51) il rilancio preda di Pirlo che poi non ci arriva # *c'è Zambrotta* # Camoranesi
(52) la palla in gioco # *c'è Schneider* # c'è Borowski però in mezzo al campo che è libero (= es. 1)
(53) adesso per Totti # capovolgimento di fronte # *c'è Pirlo* # va dentro Grosso
(54) Schneider # *c'è Materazzi* # ancora Schneider
(55) largo c'è Totti # *c'è Camoranesi* # si alza ancora la

bandierina di Hector Vergara

(56) palla recuperata da Materazzi # *c'è Totti* # ancora Camoranesi

(57) Totti # affrontato da Kehl la pressione però è giusta perché ruba palla # *c'è Odonkor* # c'è anche Friedrich

La frequenza cresce se, a queste ricorrenze, che si potrebbero considerare discorsivamente assolute, si aggiungono quelle in cui il modulo, nella sua semplicità sintattica interna, presenta tuttavia manifestazioni di collegamenti contestuali assicurate da congiunzioni o da avverbi congiuntivi, collegamenti contestuali che lo rendono parte esplicita di uno sviluppo discorsivo:

(58) il rilancio in direzione sempre di Toni # *ma c'è Camoranesi*

(59) adesso c'è largo Zambrotta che riceve ma attenzione *perché c'è Grosso* # lo ha visto

(60) da quella parte agisce Perrotta # *ma c'è anche Schneider*

(61) si è allargato Perrotta # *c'è anche Grosso*

(62) Kehl # verso Podolski # *c'è poi Gattuso*

(63) **A:** il colpo di testa di Cannavaro # un passaggio chissà come gli è venuto # Pirlo **B:** *comunque c'è sempre Cannavaro*

(64) Klose # quindi Totti di prima *ma c'è Kehl* # poi Perrotta

(65) Totti # lo costringe al passaggio non efficacissimo *ma poi c'è Kehl c'è Ballack* e si mette in movimento Podolski

(66) sul filo del fuorigioco si muove Gilardino *ma c'è anche Totti* # Zambrotta # Gattuso

Privo o dotato che sia di espliciti collegamenti contestuali, in un processo analitico che mira alla riduzione funzionale nucleare, l'interesse di tale modulo consiste nella sua immediata confrontabilità con ricorrenze semplici e assolute del nome proprio:

(67) *Totti* # *Perrotta* # la proiezione centrale

(68) con Gattuso adesso # con il colpo di Perrotta # *Pirlo* # buon inizio dei nostri quanto meno sul piano della personalità

(69) il rilancio di Zambrotta # *Camoranesi* # la rimessa laterale per la formazione tedesca

(70) arriva subito Zambrotta # *Klose* # colpo di Podolski per Ballack

(71) *Klose* # quindi Totti di prima ma c'è Kehl # *poi Perrotta*
(72) Toni gira a vuoto # *Schneider* # il pallone è verso Borowski # si guarda attorno # arriva Camoranesi # lo evita # *Lahm* # attacca ancora la Germania con Borowski
(73) stacco di Friedrich # *Pirlo* # *Podolski* # *Klose* # arriva ancora Gattuso # è dappertutto Rino # *Buffon* # *Zambrotta* # ancora per Gattuso
(74) quindi Pirlo con la collaborazione di Cannavaro # adesso Kehl con Lahm # *Ballack* # Gattuso ruba palla # *Totti* # *Gilardino* # Pirlo dà un'occhiata al movimento di Iaquinta
(75) Gattuso ha di fronte Ballack e lo supera # *Perrotta* # da questi a Grosso # *Gattuso* # in direzione ancora di Zambrotta # adesso Ballack segue Pirlo # *Totti* # affrontato da Kehl la pressione però è giusta perché ruba palla # c'è Odonkor #

Nel confronto, la variante con forma verbale finita e particella può naturalmente parere ridondante e si caratterizza di conseguenza come enfatica. Non-enfatica è, per contrasto, la menzione semplice del nome proprio, che è stata riconosciuta d'altra parte come elemento linguistico tipico delle telecronache sportive.
I nomi propri dei giocatori che intervengono nelle fasi di gioco, non necessariamente nelle più veloci e concitate, vi sono proferiti semplicemente e serialmente disposti secondo uno stile sintattico minimalmente nominale, come per es. in (67) testimonia la sequenza *Totti # Perrotta*.
Tale riduzione non si deve soltanto – o, meglio, si deve solo in minima parte – alla tendenza alla riduzione della ridondanza informativa per la presenza dell'immagine, come ha suggerito Deulofeu (2000), focalizzandone l'esistenza. Ciò che un giocatore fa si vede, sostiene lo studioso francese: dal punto di vista dell'informazione, necessario sarebbe quindi precisare solo chi è a farlo.
Il testo orale di una telecronaca è però narrativo, più che informativo, e lo è per convergenza con la narrazione per immagini, non per banale sovrapposizione. In una telecronaca, la parola non è insomma mai ridondante rispetto alle immagini. La secca nominazione dei protagonisti, nel corso della descrizione del gioco, si deve quindi piuttosto alla tendenza di questi testi a privilegiare diegeticamente i valori eponimici del nome proprio.

La nominazione vale come autenticazione narrativa di ciò che, se visto restando privo di nome proprio, è insignificante, come lo sarebbe un banale calcio ad un pallone, una banale corsa, un banale colpo di testa. Solo per via dell'associazione al nome proprio di chi lo compie, ogni atto di gioco diviene atto individuale, significativo e meritevole di menzione narrativa[9].
Come variante enfatica della secca nominazione, con cui entra peraltro da un lato in concorrenza paradigmatica, dall'altro in rapporto sintagmatico di *variatio*, il modulo *c'è Zambrotta* diluisce l'espressione: fanno così di norma le ridondanze. Riporta tuttavia apertamente l'effetto narrativo dell'agire in campo del nominato ad un valore tanto di esistenza quanto di contestuale istanza. La natura sintatticamente predicativa del *pivot* (implicita nella secca nominazione) ne viene così esplicitata[10].
Se non importa (o importa secondariamente) dire cosa faccia Zambrotta, importa affermare, nominandolo, che la nominazione non è voce di un catalogo, tanto meno puro *flatus vocis*, ma evocazione, con effetto di realtà, della pertinenza diegetica del *Dasein* di un personaggio.

6 *Esserci* e *non esserci*

Decisiva, per cogliere il valore del modulo è l'analisi per contrasto della parte pertinente del passaggio che segue, dovuta alla voce del commentatore tecnico:

[9] Lo si è argomentato in La Fauci and Necker (2010). Tali valori sono rappresentati nel modo più chiaro dalla designazione della rete che, come naturale acme agonistica dell'intero svolgimento del gioco, prende nome (e diversamente non si potrebbe) dal nome proprio del suo realizzatore: *il goal di Fabio Grosso # Mio Dio # Mio Dio #* Fabio Grosso # Fabio # Fabio Grosso # Fabio Grosso # Fabio Grosso # *tredici e trenta # tredici e trenta del secondo supplementare # Italia uno Germania zero*; *Gilardino ancora # la palla sul destro # si gira #* Alex Del Piero *# due a zero # li abbiamo annichiliti # due a zero # esulta anche Romano Prodi # due a zero #* Grosso # Alessandro Del Piero *# il timbro ancora di Alessandro # e vai Italia.*

[10] La natura sintatticamente predicativa del *pivot* (a qualunque categoria nominale esso appartenga), come la natura di puro supporto verbale delle forme di *essere* e di elemento funzionalmente pro-predicativo della particella sono state argomentate in La Fauci and Loporcaro (1995) e poste in una prospettiva romanza comparativa in La Fauci (2000: 21-40).

(76) **A:** e poi qui Zambrotta il destro # palla che si stampa sulla traversa **B:** e anche qui portiere battuto # *non c'era Lehmann*

Nel testo della telecronaca, si tratta della sola ricorrenza del modulo in esame che sia, insieme, sotto la portata predicativa d'una negazione e di un tempo grammaticale diverso dal presente.
L'imperfetto romanzo è sì tempo narrativo per eccellenza ma è anche modo dell'"Inaktualität", secondo l'opinione di Coseriu (1976: 158 e ss.). La parola del commentatore tecnico, il suo ruolo enunciativo hanno valore nel testo ovviamente in dipendenza del già avvenuto e si caratterizzano inoltre come meta-diegetici: formulano ipotesi, modulano l'effetto di realtà necessariamente introdotto dalla parola del telecronista.
Se si precisa allora il contesto di enunciazione di (76), si osserva che Mazzola proferisce *non c'era Lehmann* durante una pausa della partita, mentre sta rivedendo in video e non in presa diretta, insieme col telecronista (e naturalmente con e per i telespettatori), fasi salienti del gioco già svolto. Chi legge queste pagine riconoscerà naturalmente come familiare la situazione enunciativa che gli si sta prospettando. In una di tali fasi, il tiro di un giocatore italiano ha appunto colpito la traversa della porta avversaria e, certo, ciò non è accaduto in assenza del portiere tedesco: piuttosto in sua presenza, ma con una sua (almeno) presunta incapacità di intervenire decisivamente. Di lui Mazzola può quindi comprensibilmente affermare che "non c'era": solo come negazione di un ipotetico (*c'era*) *Lehmann*, appropriato nel caso in cui il portiere fosse intervenuto o avesse, per ipotesi, anche solo potuto farlo.
La pertinenza narrativa della nominazione del portiere tedesco consiste insomma nella valutazione di una presunta ininfluenza del suo esserci, per lo sviluppo narrativo del processo, e il valore è il medesimo che, con segno positivo, ha il modulo *c'è Zambrotta*, quando esso ricorre. E di là dalla sua apparente corrività discorsiva, proprio con questo valore, tale modulo è l'elemento funzionale nucleare intorno al quale sono poi composti i casi più complessi a diverso titolo grammaticale, esemplificati nelle pagine precedenti.
In conclusione, il modulo *c'è Zambrotta* predica (solo in apparenza pleonasticamente) la pertinenza diegetica dell'esserci del nominato e il

suo frequente ricorrere in un testo orale, come è quello di una telecronaca sportiva, in maniera al contempo idiosincratica ed emblematica, lungi dall'essere effetto della sciatteria compositiva della lingua parlata rispetto alla scritta o, peggio, alla precaria presenza nel parlato di dipendenze sistematiche, si correla rigorosamente alle funzioni espressive che la parola, l'orale non meno della scritta, manifesta con la sua stessa esistenza.

Riferimenti bibliografici

Berretta, M. (1994) Il parlato italiano contemporaneo, in M. Dardano and P. Trifone (eds) *Storia della lingua italiana*, vol. II, Torino: Einaudi, pp. 239-70.

Berruto, G. (1986) Un tratto sintattico dell'italiano parlato: il c'è presentativo, in K. Lichem, E. Mara and S. Knaller (eds) *Parallela 2. Aspetti della sintassi dell'italiano contemporaneo*, Tübingen: Gunter Narr, pp. 61-73.

Coseriu, E. (1976) *Das romanische Verbalsystem*. Tübingen: Narr.

Davies, W. and C. Rosen (1988) 'Unions as Multi-Predicate Clauses'. *Language* LXIV, pp. 52-88.

De Cesare, A.-M. (2006) *"C'è la tua bambina che gioca coi fiammiferi"*. Funzioni del costrutto presentativo c'è...che..., in A. Ferrari (ed.) *Parole frasi testi tra scritto e parlato*, «Cenobio», 3 (2006), pp. 215-21.

Deulofeu, H.-J. (2000) Les commentaires sportifs télévisés sont-ils un genre au sens de la «grammaire des genres»?, available on-line from http://jose.deulofeu.free.fr/article.php3?id_article=12 (accessed: 30 settembre 2009).

Jakobson, R. (1973 [1966]) Le parallélisme grammatical et ses aspects russes, in *Questions de poétique*, Paris: Seuil, pp. 234-79 ['Grammatical parallelism and its Russian facet'. *Language* XLII, 2, pp. 399-429].

La Fauci, N. (2000) *Forme romanze della funzione predicativa. Teorie, testi, tassonomie*. Pisa: ETS.

La Fauci, N. and M. Loporcaro (1995) 'Outline of a theory of existentials on evidence from Romance'. *Studi italiani di linguistica teorica e applicata* XXVI, pp. 5-55.

La Fauci, N. and H. Necker (2010) Nomi propri (e fatti correlati) in

una telecronaca sportiva, in N. Pepin and E. De Stefani (eds) *Eigennamen in der gesprochenen Sprache,* Tübingen: Francke, pp. 200-223.

Necker, H., S. Simon and L. Tronci (2009) Paratassi e ipotassi in una telecronaca sportiva, in A. Ferrari (ed.) *Sintassi storica e sincronica dell'italiano. Subordinazione, coordinazione, giustapposizione. Atti del X Congresso della Società Internazionale di Linguistica e Filologia Italiana (Basilea, 30 giugno-3 luglio 2008)*, Firenze: Franco Cesati, pp. 1375-1391.

Sabatini, F. (1980) Linee di tendenza dell'italiano contemporaneo e problemi di norma, in L. Lindgren (ed.) *La lingua italiana in Finlandia*, Turku: Sezione di Italiano, Facoltà di Lettere, Università di Turku, pp. 73-91.

Salvi, G. (1988) La frase semplice, in L. Renzi, G. Salvi and A. Cardinaletti (eds) *Grande grammatica italiana di consultazione*, 3 voll., Bologna: il Mulino, vol. I, pp. 37-127.

Schiller, A. (1995) Präsentativa im Französischen und Italienischen. Einige syntaktisch-vergleichende Überlegungen, in C. Schmitt and W. Schweickard (eds) *Die Romanischen Sprachen im Vergleich*, Bonn: Romanistischer Verlag, pp. 289-304.

Venier, F. (2002) *La presentatività. Sulle tracce di una nozione*. Alessandria: Edizioni dell'Orso.

Le espressioni di sentimento nel LIP

Alberto Maria Langella, Simona Messina, Serena Santonicola
Università degli Studi di Salerno
alangella@unisa.it; smessina@unisa.it; ssantonicola@unisa.it

Abstract

This contribution is aimed to test and classify a restricted number of sentiment expressions which: paura, spavento, preoccupazione; contentezza, allegria, divertimento; meraviglia, entusiasmo, interesse. We suggest a description of syntactic structures where the psychological expressions occur, starting from the analysis of different sentence typologies retrieved by analysing and screening the whole Corpus LIP (De Mauro et al*, 1993).*
According to the lexicon-grammar approach, verbs, nouns and adjectives are considered as operators (Harris, 1976, 1982) that express either feelings or causative agent of sentiment, and whose first argument is Nhum, a human subject who feels the emotion analyzed.
Thanks to the main tools of computational linguistics and using Silberstein's software NooJ, our research provides also an automatic and formal recognition (so called graph) of the sentiment expression, concerning the items we have previously chosen. With these graphs intended to underline the difference between the sentence structures concretely used in spoken language and the sentence structure potentially correct and grammatically acceptable.
Far from claiming of being a completely exhausted study, our survey introduces also the question of support verb and their modal, aspectual and stylistic variations.

0 Introduzione

Questo studio[1] intende coniugare la prospettiva teorica del Lessico-Grammatica, con particolare riferimento alla «grammatica locale dei sentimenti» (Gross, 1995: 105), con un'applicazione su un campione rappresentativo di testi quale il corpus del *Lessico di frequenza dell'italiano parlato* (LIP; De Mauro *et al*, 1993). Tale applicazione

[1] Questa ricerca, frutto di un intenso lavoro di gruppo, può essere divisa in tre sezioni. La prima, dedicata agli approfondimenti teorici e alla definizione del quadro metodologico generale è stata curata dalla dott.ssa Santonicola. La seconda sezione, concernente la classificazione delle strutture di frasi e l'analisi dei dati del LIP è stata realizzata dalla dott.ssa Messina. Infine il riconoscimento automatico dell'espressioni di sentimento è stato operato dal dott. Langella.

ha l'obiettivo di classificare alcuni "predicati di sentimento" relativi a tre microclassi:
1. *paura*, *spavento*, *preoccupazione*; 2. *contentezza*, *allegria*, *divertimento*; 3. *meraviglia*, *entusiasmo*, *interesse*, che si realizzano nelle diverse tipologie di frasi a seconda che l'operatore sia verbale, nominale, aggettivale o avverbiale.
La descrizione automatica e formale di dette frasi riguarda due diverse strutture che, secondo la grammatica in "operatori ed argomenti", possono essere così sintetizzate:

• frasi con struttura On, in cui l'operatore è l'espressione di sentimento nelle sue diverse forme e l'argomento è *Num*, soggetto che prova il sentimento:
es: *Lia è allegra*; *Sofia ha paura*; *Eva si diverte*;
• frasi con struttura Ooo, in cui interviene l'agente o la causa del sentimento (in forma di gruppo nominale o di frase), collegato a *Num*, soggetto che prova il sentimento, attraverso un operatore causativo del tipo *fare*, *causare*, *essere causa*, ecc.:
es: *Questa situazione ha causato l'allegria di Lea*; *Uscire di notte fa paura a Max.*

1 L'approccio lessico-grammaticale

La nostra analisi è una diretta applicazione della teoria del Lessico-Grammatica, che individua gli elementi di base di significato non nelle singole parole autonome, ma in famiglie di frasi (Gross, 1975). Il principio fondamentale è che l'unità minima del lessico è rappresentata dalle frasi elementari e non dalle parole.
La componente teorica principale del Lessico-Grammatica è rappresentata dalla "Grammatica in Operatori e Argomenti" di Z.S. Harris. Harris considera il linguaggio come un sistema le cui proprietà sono quelle di un ente matematico, cioè di un oggetto i cui elementi sono in rapporto di reciproca dipendenza. La grammatica harrisiana parte dall'analisi di "concatenazioni" di stringhe di caratteri: un discorso è una concatenazione di sequenze organizzata a partire dalle relazioni esistenti tra gli operatori (o) e i loro argomenti (n) e tra gli operatori stessi. Ogni operatore sarà, quindi, definibile come una funzione e i suoi eventuali argomenti come le variabili di questa funzione. In tal senso, un operatore è individuabile come un elemento di connessione tra gli argomenti del discorso (o della

frase). L'appartenenza a classi di parole diverse non impedisce che lemmi differenti, ma morfologicamente o foneticamente correlati tra loro, possano ugualmente svolgere la funzione di operatore. Se le frasi possono essere interpretate come varianti morfo-fonetiche nelle quali la relazione stabilita tra l'operatore e i propri argomenti resta costante e inalterata, si dice che tali elementi di frase intrattengono tra loro correlazioni trasformazionali (per es., una frase con operatore verbale può essere in correlazione con una frase nominale o aggettivale).
Tenendo conto che un "sentimento" può realizzarsi con operatori verbali, nominali e aggettivali, l'obiettivo generale della nostra indagine è, quindi, la determinazione dell'insieme di frasi possibili che possono avere, in equivalenza distribuzionale e parafrastica, forma verbale, aggettivale e nominale:

Nello s'impaurisce ≡ Nello è impaurito ≡ Nello ha paura.

Nella nominalizzazione e nell'aggettivalizzazione intervengono i verbi supporto come *avere*, *essere* oppure le loro varianti aspettuali, modali e stilistiche:

Nello è impaurito → Nello resta impaurito
Nello ha paura → Nello prova paura.

2 L'analisi lessico-grammaticale del LIP

Come sostenuto in D'Agostino (1993) l'analisi lessicale del parlato di una lingua esplica gli usi sintattico-semantici delle varie forme con maggiore informatività, in quanto le classificazioni lessico-grammaticali, applicate a corpora testuali significativi come il corpus LIP, descrivono le restrizioni che influenzano l'equiprobabilità di occorrenze delle parole l'una rispetto all'altra e quindi definiscono la struttura predicativa di registri e usi diversi di una lingua. Il parlato, pur se teoricamente vincolato ad alcune possibili regole previste dal sistema lingua, è fortemente condizionato dalla situazione di enunciazione, dalla relazione sociale tra i parlanti e dalle contingenze di realizzazione. I fenomeni del parlato, pertanto, non possono essere esaustivamente studiati se non avvalendosi di dati empiricamente osservabili e, quindi, sistematicamente registrabili.

Esaminando nel dettaglio le diverse forme identificabili nel LIP si registra il quadro di alcune possibilità del sistema che sono concretamente realizzate.
All'interno del vasto panorama testuale offerto dal corpus LIP, si è inteso analizzare la distribuzione e la tipologia di alcuni operatori di sentimento e quindi la distribuzione e la tipologia di verbi supporto ed estensioni di verbo supporto in caso di nominalizzazione, aggettivalizzazione e costruzione causativa.

3 Le grammatiche formali

I principali strumenti d'analisi della linguistica computazionale sono i meccanismi di riconoscimento detti *parser*. Essi si affidano a strutture matematiche dette *grammatiche formali*. Tali strumenti computazionali danno anche luogo a descrizioni e misurazioni accurate di frequenze, variazioni di frequenze, fino alla descrizione sperimentale di diversi fenomeni strutturali, come la definizione di un quadro delle regolarità e delle variazioni d'uso delle entrate lessicali.
Il nostro lavoro è costituito dall'elaborazione di automi finiti su base categoriale, i quali rappresentano un sottoinsieme delle diverse grammatiche formali possibili. In un primo momento vengono descritte le diverse strutture possibili in cui l'espressione di sentimento (di seguito *Esent*) può essere nella posizione sintattica nominale (N_0, N_1, N_2), verbale (*V*) o aggettivale (*Agg*). In un secondo momento si rintracciano nel corpus LIP le concrete realizzazioni frastiche che si producono nel parlato quando si esprimono i sentimenti. Di esse si darà un descrizione tramite i grafi, da leggersi da sinistra a destra, dallo stato iniziale allo stato finale.

4 La selezione dei lemmi

Le microclassi di lemmi prese in esame riguardano tre gruppi di *V-n* in relazione morfo-fonologica con i corrispondenti verbi a oggetto diretto con una completiva soggetto e con struttura: *Ch F V* N_1 (Elia, 1984: 85):

paura → impaurire; spavento→ spaventare; preoccupazione → preoccupare; contentezza → contentare/accontentare; allegria →

rallegrare; divertimento → divertire; meraviglia → meravigliare; entusiasmo → entusiasmare; interesse → interessare.

Questa classe ha una certa coerenza semantica poiché la maggioranza dei verbi ha un valore "psicologico" o, più esattamente, si tratta di verbi relativi ad un sentimento che un soggetto agente produce in un soggetto umano in posizione N_1. L'approccio teorico si basa sulle classificazioni lessico-grammaticali dei verbi dell'italiano (EMDA, 1981; Elia, 1984; Vietri, 2004) e sulle più recenti argomentazioni teoriche proposte da D'Agostino (2007) che, relativamente alle completive, affronta la questione partendo dalla frase di base ad operatore nominale dalla quale è possibile ricostruire un percorso di equivalenze frastiche del tipo On.

5 Frase nominale con struttura On

I verbi della classe citata hanno un comportamento omogeneo. N_0 può essere realizzato con stringhe distribuzionalmente equivalenti (*Num*, *Nnr*, *il fatto Ch F*, V^0 *inf*). *V* è rappresentato dal verbo di sentimento (es. *impaurisce* + *spaventa* + *preoccupa* + *accontenta* + *rallegra* + *diverte* + *meraviglia* + *entusiasma* + *interessa*). I nomi operatori, invece, presentano un numero più o meno ampio di restrizioni in relazione all'uso dei verbi supporto, del determinante e del modificatore.
Innanzi tutto bisogna distinguere i verbi supporto come *avere*, *essere* e *essere in* dalle loro varianti aspettuali, modali e stilistiche. A*vere* non si combina con tutti i nomi da noi isolati, e anche laddove può essere usato, spesso richiede la variazione di categoria grammaticale singolare/plurale, l'introduzione di un determinante o di modificatore o anche di un classificatore composto del tipo *una sensazione di*, *un'espressione di*:

1) *Nello ha paura*
2) *Nello ha (*preoccupazione + preoccupazioni)*
3) *Nello ha una grande preoccupazione.*

Con le varianti *provare*, *rivelare* è possibile sostituire a *Num* un *Npc* adeguato in relazione metonimica, ottenendo frasi accettabili:

4) *Gli occhi di Nello rivelavano allegria*
5) *Il sorriso di Nello rivela un grande divertimento.*

5.1 Frase aggettivale con struttura On

L'applicazione dell'operatore aggettivale corrispondente nella struttura On si realizza con frasi a verbo supporto, del tipo:

6) *Lia è (contenta + allegra)*

dove il supporto *essere* può comparire nelle sue varianti aspettuali:

7) *Lia (appare + sembra + pare + risulta + diventa) (allegra + impaurita).*

5.2 Frase verbale con struttura On

La struttura ad operatore verbale On si costruisce con la corrispondente forma intransitiva autonoma:

8) *Lia s'impaurisce*
9) *Lia si diverte*
10) *Lia si meraviglia.*

Come sostenuto in D'Agostino (1992:176):

è necessario parlare di forma intransitiva autonoma, dato che per la forma 'si' non si riesce a rintracciare nessun antecedente nominale come nel caso delle vere proforme.

Sono, infatti, inaccettabili le frasi:

8a) **Lia impaurisce se stessa*
9b) **Lia diverte se stessa*
10c)**Lia meraviglia se stessa.*

6 La struttura Ooo

Con l'introduzione del soggetto causativo si aggiunge alla struttura On un altro operatore costituito dagli equivalenti distribuzionali (*Num + Nnr +il fatto Ch F +V^0 inf*) che possono essere ridotti in (*Num + F*) (D'Agostino, 2007: 353). Le frasi elementari nominali, aggettivali e verbali si trasformano così da On a Ono:

11) *Nello ha paura del temporale*
12) *Lia è meravigliata dell'esito delle elezioni*
13) *Nello si preoccupa del fatto che la figlia parta.*

Le forme nominali, poste a destra della frase semplice esprimono la causa del sentimento, per cui il percorso può procedere con l'inserimento dell'operatore causativo Ooo del tipo *causa/causare*:

11a) *Il temporale è la causa del fatto che Nello abbia paura*
12a) *L'esito delle elezioni è la causa del fatto che Lia sia meravigliata*
13a) *Che la figlia parta è la causa del fatto che Nello si preoccupa.*

L'operatore causativo può essere inoltre sostituito con gli equivalenti: *provocare*, *suscitare* o con gli aspettuali *alimentare*, *diminuire*, *accrescere* ecc., mentre la frase aggettivale si può costruire anche con i causativi *rendere* o *fare*:

14) *Il buio rende impaurito Nello*
15) *Lia fa (E+ che sia) contento Nello.*

L'ultima realizzazione è la costruzione a completiva in posizione soggetto con riduzione ad infinitiva in caso di coreferenza dei soggetti:

16) *Prendere l'aereo spaventa Nello*

che si associa all'altra costruzione aggettivale, nelle *Esent* che lo consentono:

17) *Prendere l'aereo è (pauroso + spaventoso + preoccupante + divertente + meraviglioso + entusiasmante + interessante) per Nello*.

7 I dati dal LIP

Gli esempi fin qui esposti sono ipotesi di laboratorio, teoricamente possibili ma non facilmente riscontrabili, nella loro completezza, nella lingua parlata che, per sua natura, è frammentaria e non sempre coerente. Il discorso parlato si sviluppa attraverso una serie di fratture dovute a cancellazioni, ripensamenti della strategia discorsiva, interruzioni e silenzi, per cui, mentre è facile isolare i

singoli lemmi, non lo è per quanto riguarda le frasi che molto spesso sono interrotte e si completano in enunciati successivi. L'analista deve perciò cercare di ricostruire, laddove sia possibile, la logica degli enunciati, senza tuttavia nulla aggiungere a quanto detto dal parlante.
Tenuto conto di tali problematiche si è giunti alla compilazione delle tavole sintetiche che sono state divise secondo i seguenti criteri:

Frase nominale a verbo supporto (1)→ frase a strutture On/Ono, sia con verbo supporto neutro sia con i suoi sostitutivi aspettuali e stilistici.
Frase nominale ad operatore causativo (2)→ frase a struttura Ooo con l'operatore causativo *causare* o i suoi sostitutivi.
Altri usi del nome (3)→ uso dei nomi di sentimento non collocabile nelle due categorie precedenti.
Frase aggettivale a verbo supporto con aggettivo in -to/0 (4)→ frase in cui l'aggettivo si riferisce al soggetto esperiente che subisce il sentimento. L'aggettivo può avere suffisso *-to* o essere a suffisso 0.
Frase aggettivale a verbo supporto con aggettivo in -oso/-nte (5)→ frase in cui l'aggettivo si riferisce al soggetto agente che provoca il sentimento nel soggetto esperiente che può essere espresso o sottinteso così come accade nell'uso assoluto di alcuni *Vsent*: *Il mare di notte spaventa – Il mare di notte è spaventoso.* Frasi di questo tipo hanno una struttura Oo[2] perché equivalenti a *Il mare di notte (fa + causa + suscita + ecc.) spavento.*
Altri usi dell'aggettivo (6)→ uso degli aggettivi di sentimento non collocabile nelle due categorie precedenti.
Frase ad operatore verbale (7)→ frase con verbo operatore di sentimento sia nella forma transitiva (struttura Oo/Ooo) sia in quella intransitiva autonoma (On/Ono).

I 699 casi esaminati sono riportati in tre tavole:

Casi esaminati = 204

Sentimento	**1**	**2**	**3**	**TOT.** *Nsent*	**4**	**5**	**6**	**TOT.** *Asent*	**7**
paura	43	8	18	**69**	-	-	1	**1**	**2**
spavento	-	-	-	-	3	1	8	**12**	**8**
preoccupazione	14	1	12	**27**	7	2	3	**12**	**73**

Tavola 1: paura – spavento – preoccupazione

[2] I 2/3 dei verbi di sentimento ammettono un uso assoluto, cioè una sottostruttura N_0V con N_1 cancellato (cfr. Elia, 1984: 86).

Casi esaminati = 153

Sentimento	**1**	**2**	**3**	**TOT.** *Nsent*	**4**	**5**	**6**	**TOT.** *Asent*	**7**
contentezza	-	-	1	**1**	45	-	13	**58**	**7**
allegria	1	-	5	**6**	2	-	2	**4**	**-**
divertimento	1	-	4	**5**	-	16	14	**30**	**42**

Tavola 2: contentezza – allegria – divertimento

Casi esaminati = 342

Sentimento	**1**	**2**	**3**	**TOT.** *Nsent*	**4**	**5**	**6**	**TOT.** *Asent*	**7**
meraviglia	-	-	9	**9**	2	10	23	**35**	**9**
entusiasmo	3	1	4	**8**	-	-	1	**1**	**4**
interesse	41	2	48	**91**	4	27	35	**66**	**119**

Tavola 3: meraviglia – entusiasmo – interesse

Sentimento	**Totale Esent**	**Usi nominali**	**Usi aggettivali**	**Usi verbali**
paura	72	**96%**	1,3%	2,7%
spavento	20	-	**60%**	40%
preoccupazione	112	24%	10,8%	**65,17**
contentezza	66	1,5%	**87,8%**	10,6%
allegria	10	**60%**	40%	-
divertimento	77	6,5%	40%	**54.5%**
meraviglia	53	17%	**66%**	17%
entusiasmo	13	**61,5%**	7,7%	30,8%
interesse	276	**34**	24%	43%

Tavola 4: valore percentuale delle occorrenze

Nonostante la loro esiguità i numeri ci suggeriscono alcune considerazioni.
Innanzitutto si è potuto notare che le *Esent* hanno comportamenti distribuzionali diversi.
Per quanto riguarda gli usi nominali, i risultati più significativi riguardano indubbiamente il lemma *paura*: su 72 usi totali 69 sono nominali e, di questi, 41 sono in frasi a *Vsup*; i restanti 28 sono variamente classificati. La frase nominale è per massima parte costruita con i *Vsup avere*; talvolta *avere* è sostituito da *mettersi paura*, che ha un valore legato al campione regionale napoletano:

FB16	C: di questa cosa il Barbetti **ha paura** io lo so che lui **ha paura**

NE5	A: aah e diglielo dai microfoni **ti metti paura di Assunta**

Il verbo *preoccupare*, nel campione esaminato, è usato soprattutto nelle formule di cortesia:

RB9	H: non si non si non **non si preoccupi** io domani mattina verifico

Nelle frasi nominali a verbo supporto è frequente l'uso di *manifestare, esprimere* come varianti di *avere*.

MC2	A: In questo incontro abbiamo **manifestato le preoccupazioni nostre**

Contentezza, *divertimento* e *meraviglia* vengono espressi preferibilmente con le frasi aggettivali:

FD17	A: …anche laggiù facevano la festa del partito bambini che **rimangono meravigliati** a sentire per la prima volta eh un'apparecchio eh radio… loro **si son trovati meravigliati** nel sentire la loro voce

In conclusione, questa breve analisi di espressioni di sentimento conferma quanto già sostenuto da Gross come riportato in Vietri (2008: 382):

> […] le ricerche di Gross hanno mostrato che i comportamenti sintattici, indissolubili dai concreti elementi lessicali, sono al limite dell'individualità: ogni nome, ogni verbo, ogni aggettivo, ogni avverbio ogni determinante ha una sua propria caratterizzazione individuale, che rientra solo superficialmente all'interno di classi di regole più ampie.

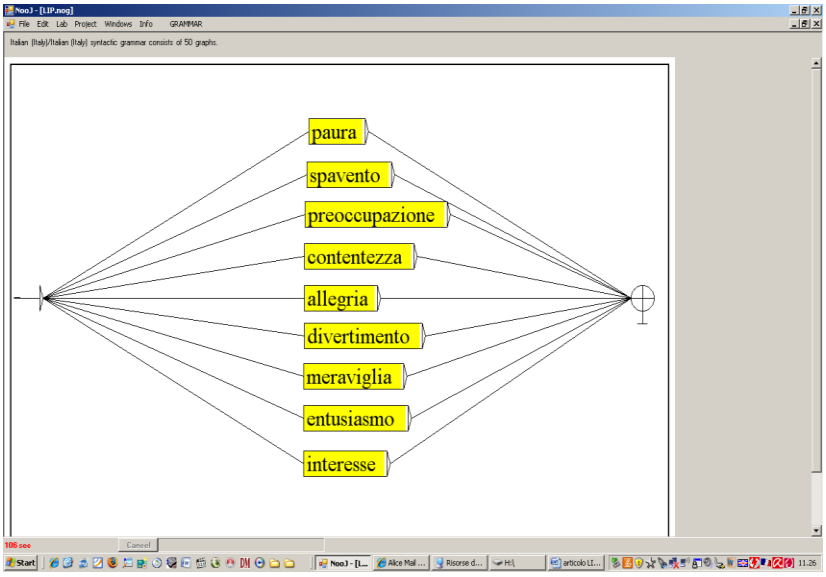
8 Il riconoscimento automatico delle espressioni di sentimento nel LIP

8.1 Il programma utilizzato

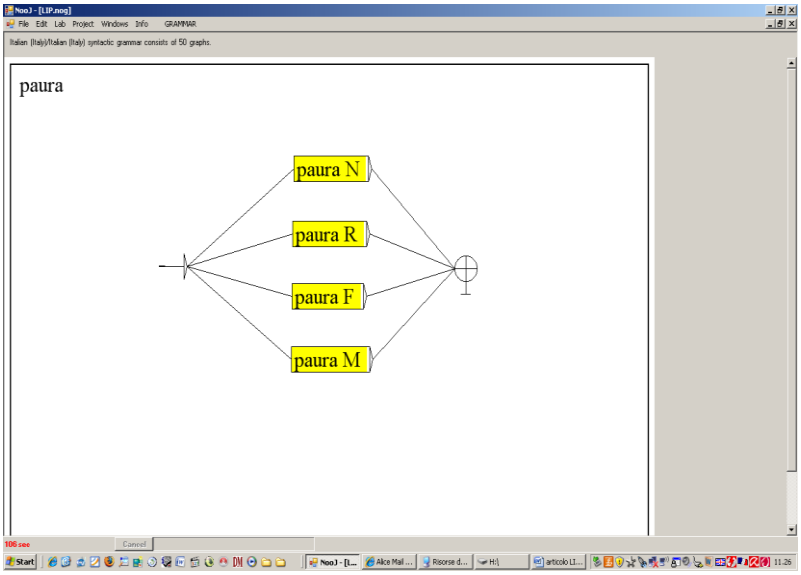
Per il parsing delle occorrenze relative alle espressioni di sentimento ci siamo serviti di NooJ, un programma ideato da Max Silberztein ed in uso presso l'Università di Besançon (Francia). Questo software è l'evoluzione del programma INTEX, utilizzato, a partire dal 1993, da Maurice Gross presso il LADL dell'Università Parigi VII. Il modulo italiano (NooJ IT) è stato sviluppato presso il Dipartimento di Scienze della Comunicazione dell'Università degli studi di Salerno da Simonetta Vietri.
Il programma è basato sulle due principali risorse linguistiche costituite da Dizionari elettronici di grossa taglia e da grafi o automi a stati finiti utilizzati come risorsa morfologica e di riconoscimento di fenomeni sintattici.

8.2 L'automa di riconoscimento delle espressioni di sentimento nel LIP

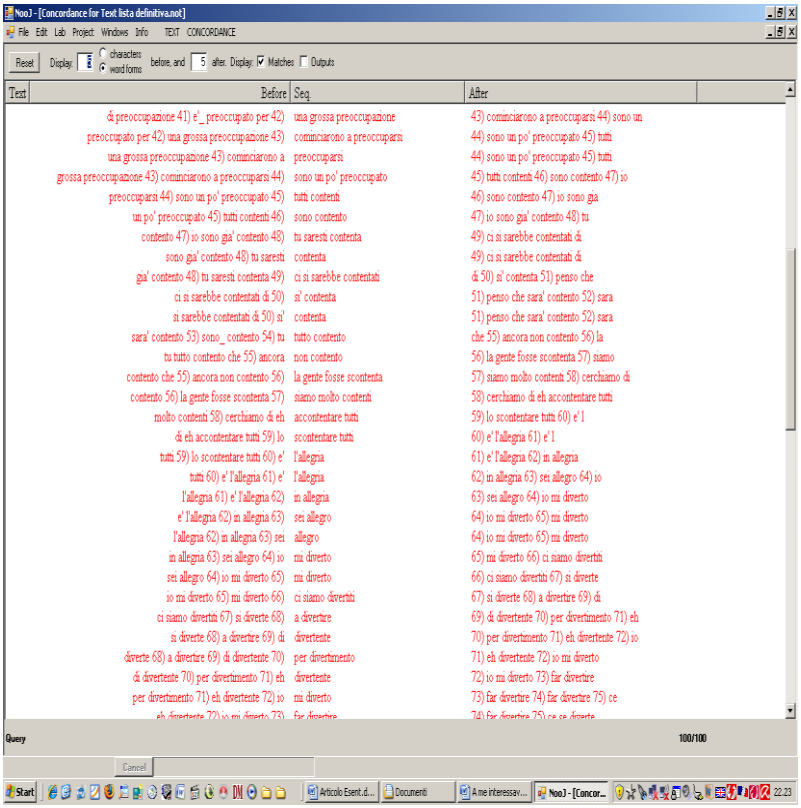
La metodologia che abbiamo seguito è stata quella di costruire un meccanismo di riconoscimento delle espressioni di sentimento nel LIP Partendo dai nomi selezionati, ci siamo poi occupati del riconoscimento di occorrenze verbali ed aggettivali morfo-fonologicamene imparentate (*impaurirsi*, *spaventarsi*, *interessante*, *allegro* ecc.).
L'automa prodotto è il seguente:



Ed all'interno di ciascuno dei nodi di colore giallo ve ne sono altri quattro:



Ciascun dei quattro nodi (Paura N, Paura R, Paura F, Paura M) contiene i grafi di riconoscimento di tutte le occorrenze della costellazione parafrastica costruita a partire da *paura* all'interno del LIP Per le altre 8 costellazioni si è proceduto nello stesso modo.
Il totale è di 36 automi (4 città per 9 nomi) racchiusi all'interno di un unico grafo denominato LIP.nog. Il grafo LIP.nog è un meccanismo di riconoscimento costruito *ad hoc* per il trattamento automatico delle espressioni di sentimento nel LIP. Il suo parser sull'intero corpus dell'italiano parlato produce alcune centinaia di occorrenze riconosciute automaticamente. Forniamo un estratto esemplificativo del suo funzionamento sulle prime 100 occorrenze rintracciate:



8.3 Il riconoscimento delle forme dialettali

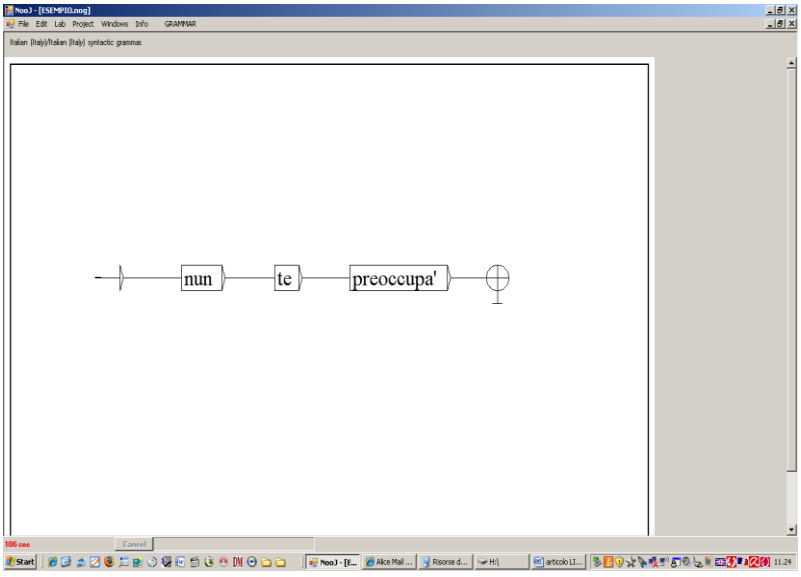
Molto spesso nel nostro tentativi di descrivere e riconoscere computazionalmente le espressioni di sentimento nel LIP, ci siamo trovati dinanzi al problema riguardante le numerose espressioni linguistiche dialettali. Es.: *Nun te preoccupa'*; *Un te preoccupa'*.
Mentre abbiamo ritenuto opportuno effettuare la normalizzazione ortografica del testo[3], non abbiamo voluto procedere nello stesso modo per le forme dialettali del tipo mostrato poc'anzi. Il motivo di questa scelta è che un processo di normalizzazione di questi tipi

[3] Per poter procedere all'analisi automatica è stato necessario normalizzare il corpus LIP. La normalizzazione ha riguardato solamente: la trasformazione degli apostrofi impropri in accenti (es. *c'e'* → *c'è*; *puo'* → *può*; *pero'* → *però*); l'eliminazione dei seguenti simboli: #, ##, ### = pausa breve, media e lunga; sta<te> = parola ricostruita; -pe- = parola non ricostruibile; ciao_ = tenuta vocalica in fine di parola.

avrebbe tradito completamente la natura e le caratteristiche testuali del LIP, e cioè un corpus del parlato, all'interno del quale le forme dialettali devono trovare una loro legittima collocazione.
Procedere in questo modo ci ha indubbiamente creato alcuni problemi di carattere metodologico. Infatti il funzionamento del parser da noi utilizzato, e cioè Nooj, si basa su informazioni grammaticali fornite al programma sotto forma di dizionari in formato elettronico. La nostra scelta quindi, di descrivere le *Esent* così come rintracciate nel testo, e cioè in forme spesso dialettali, ci ha impedito di poter usare le informazioni linguistiche presenti nei dizionari del programma, originariamente concepito solo ed esclusivamente per le forme standard dell'italiano.
Ciò comporta che per la frase seguente

RB21	A: nun te preoccupa' de questioni amministrative

non si è potuto procedere al riconoscimento automatico della sequenza avvalendosi delle informazioni linguistiche dei dizionari, ma si è provveduto a specificare una ad una le stringhe in esame:

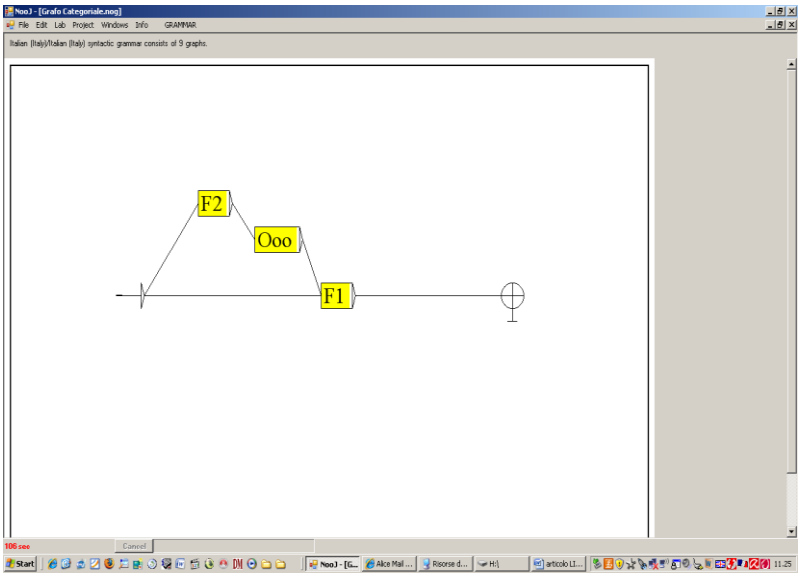


Ciò ci ha senz'altro creato problemi di carattere metodologico, ma nello stesso tempo ha fornito alla nostra descrizione linguistica un

maggiore rispetto delle peculiarità linguistiche di un corpus di parlato come il LIP.

8.4 L'automa su base categoriale

Si procede ora all'elaborazione di un automa su base categoriale (Grafo Categoriale.nog) finalizzato alla descrizione delle stesse *Esent* studiate all'interno del LIP, ossia un automa tarato più sulla "competenza" che sugli "usi" concreti della lingua parlata. Quindi, non scendiamo nella descrizione dettagliata delle specifiche occorrenze come fatto per il LIP, ma teniamo la nostra descrizione al livello delle diverse categorie grammaticali.
Il confronto tra Grafo Categoriale.nog e LIP.nog ci fornisce la possibilità di paragonare la tipologia di frasi concretamente realizzate nel corpus dell'italiano parlato e la tipologia più vasta di frasi potenziali sotto il vincolo della dipendenza dagli stessi operatori *Esent* studiati all'interno del nostro articolo.
L'automa ha la struttura seguente:



che riconoscerà frasi come le seguenti:

18) *Mario è preoccupato*
19) *Maria ha paura*

20) *Il fatto che Mario parta preoccupa Maria*
21) *La venuta di Mario spaventa Maria.*

Le differenze forniteci dai due modelli descrittivi rappresentati dai grafi LIP.nog e Grafo Categoriale.nog, con il primo dedicato alla descrizione degli operatori discussi in questo articolo nel LIP, ed il secondo dedicato al riconoscimento di frasi teoriche tutte circoscritte nel perimetro della competenza linguistica, forniscono chiare indicazioni relative all'enorme differenza tipologica tra la lingua scritta e quella parlata.
La maggior parte delle frasi riconosciute da Grafo Categoriale.nog sono assenti nel LIP:

22) (Num_y + F_y) *è la causa del fatto che* Num_x *si preoccupi*
23) (Num_y + F_y) *causa il fatto che* Num_x *si impaurisca*
24) (Num_y + F_y) *è la causa del fatto che* Num_x *si spaventi*
25) (Num_y + F_y) *causa il fatto che* Num_x *si rallegri*
26) (Num_y + F_y) *è la causa del fatto che* Num_x *si meravigli*
27) (Num_y + F_y*) causa il fatto che* Num_x *si entusiasmi.*

Si tratta di frasi in cui Num_y sta per un agente esterno e Num_x per il soggetto umano che "prova" il sentimento; F_y invece indica la possibilità che in quella posizione possa occorrere un frase.
In alcuni casi nel LIP forse è possibile rintracciare delle frasi che sembrano rientrare nei tipi mostrati ora, ma si tratta quasi sempre di tracce di strutture più lunghe non manifeste in superficie all'interno del discorso e difficilmente deducibile dalla lettura dei corpora.

9 Conclusioni

La categoria "espressioni di sentimento" pone il problema della sua delimitazione globale: se si può, per sgomberare il terreno, ipotizzare un insieme di criteri distribuzionali o sintattici, questi, verificati su di un corpus esteso non appaiono né specifici né tanto meno generali. Ciascuna espressione, infatti, definisce un paradigma particolare di cui non è sempre facile scoprire il principio di costituzione. Questo vale soprattutto per le costruzioni nominali, che presentano, come di sopra accennato, una serie di restrizioni relative all'uso dei verbi supporto, dei determinanti e dei modificatori.

Con questo lavoro abbiamo inteso innanzitutto approfondire le diverse combinazioni *Vsup Nsent*, per cui un primo catalogo di difficoltà da affrontare è la selezione dei verbi supporto, tenendo conto che essi offrono un gran numero di combinazioni che prevedono sia l'uso dei *Vsup* semplici, ossia di quei verbi come *avere* ed *essere* che sembrano non avere altro ruolo se non quello di ristrutturare la frase senza cambiarne il senso, alle loro estensioni che *ne modifient ni la distribution du sujet ni le régime que l'on observe avec le Vsup simples* (Gross, 1981: 33), a *Vsup* contraddistinti da valori aspettuali, stilistici ed eventualmente regionali, che apportano degli elementi di senso più marcati.

Riferimenti bibliografici

D'Agostino, E. (1992) *L'analisi del discorso*. Napoli: Loffredo.

– (1993) *Sociolinguistica computazionale. Un'applicazione descrittiva al corpus del LIP* Salerno: Ed. 10/17.

– (2007) Classificazioni lessico-grammaticali: ovvero, liberarsi dall'angoscia, in A. Elia and A. Landi (eds) *Quaderni del Dipartimento di Scienze della Comunicazione dell'Università degli Studi di Salerno*, Roma, Carocci, pp. 343-58.

De Mauro, T., F. Mancini, M. Vedovelli and M. Voghera (1993) *Lessico di frequenza dell'italiano parlato*, Fondazione IBM Italia. Milano: Etas Libri.

EMDA, Elia, A., M. Martinelli and E. D'Agostino (1981) *Lessico e strutture sintattiche.* Napoli: Liguori.

Elia, A. (1984) *Le verbe italien.* Paris: Nizet.

Gross, M. (1975) *Méthodes en syntaxe*. Paris: Hermann.

– (1981) 'Les bases empiriques de la notion de prédicat sémantique'. *Langages* 63, pp.7-52.

– (1995) 'Une grammaire locale de l'expression des sentiments'. *Langue Française* 105, pp. 70- 87.

Harris, Z.S. (1976) *Notes du cours de syntaxe*. Paris: Le Seuil.

– (1982) *A Grammar of English on Mathematical Principles*. New York: John Wiley&Sons.

Vietri, S. (2004) *Lessico-grammatica della lingua italiana. Metodi, descrizioni e applicazioni*. Torino: UTET.

– (2008) 'La grammatica dei determinanti'. *SILTA* 2, 2008, pp. 381-401, Roma: Pacini Editore.

L'immediatezza comunicativa nelle lingue romanze.
Figure di sintassi nel C-ORAL-ROM

Araceli López Serena
Universidad de Sevilla (Spagna)
cheilop@us.es

Abstract

In relation to the study of what has been called the language of immediacy (to which we can also refer as communicative closeness or proximity) (Koch and Oesterreicher 1985; 1990/2007; 2001), the field of spoken syntax has deserved little or inadequate attention until the date. Its analysis has been mostly done from a written biased point of view, colloquial syntactic structures being considered, in the best cases, as irregularities with respect to normative or grammatical structures, if not simply as grammatical mistakes or agrammatical occurrences. In order to avoid this written bias, I have turned to the analytical tools developed by the Groupe Aixois de Recherche en Syntaxe (GARS). What Blanche-Benveniste called syntax figures make us able to give account of colloquial syntax without having to make reference to written structures. These figures allow the analysis of the traces left by the discourse production process under circumstances of maximal communicative immediacy, that are always eliminated when editing written texts, but remain in oral transcriptions. In all cases we have to do with supraindividual patterns universally employed in the construction of unplanned discourse, so my main purpose will be to show this kind of patterns in Italian, where to my knowledge this sort of analysis has not yet been forehanded. My data will be extracted mainly from Italian transcriptions from C-ORAL ROM (Integrated Reference Corpora for Spoken Romance Languages), but I will also use Spanish examples from the same corpus in order to show that we are dealing with universal patterns that have to be expected in any language. I will document in this corpus the different formal types of syntax figures I have distinguished, and I will complete the analysis with some functional considerations in order to better understand to what communicative purposes correlate the different universal construction patterns that can be observed in immediate discourses.

1 Introduzione[*]

Lo studio dell'immediatezza comunicativa deve oggi affrontare numerose questioni problematiche. Da una parte, esistono problemi di imprecisione terminologica e indeterminatezza concettuale, che non sono stati ancora completamente superati. Infatti gli studiosi della materia usano in modo indiscriminato termini come linguaggio, lingua e discorso (coniati da Coseriu per i livelli universale, storico e attuale dello studio linguistico); fatto questo che si ripresenta anche nel caso di termini come parlato, orale, colloquiale, popolare, volgare ecc. (cfr. López Serena, 2007a). Dall'altra, sia i parametri situazionali proposti per caratterizzare *esternamente* queste varietà che i tratti linguistici suggeriti per una sua caratterizzazione *interna* sono provvisori e hanno bisogno di analisi approfondite che permettano di delimitare con chiarezza le differenti varietà concettuali (vale a dire, le differenze fra un uso della lingua – *socialmente, fisicamente, tematicamente,* ecc. (cfr. Koch/Oesterreicher, 1985; 1990/2007; 2001) – immediato o più informale e un uso distante o più formale). Comunque, anche se le difficoltà riguardano tutti i livelli di analisi, la sintassi è sempre stata il vero tallone di Achille.

Come è noto, tra le insufficienze già segnalate nella descrizione della sintassi colloquiale spicca il fatto che tradizionalmente essa sia stata studiata da una prospetttiva scritturale. Ne consegue che le strutture sintattiche colloquiali sono state considerate, nel migliore dei casi, come irregolarità (anacoluti, spostamenti, ellissi, concordanze mancate ecc.) rispetto alle strutture normative, quando non semplicemente degli "errori grammaticali". Quindi la terminologia sintattica tradizionale è tutt'altro che adeguata ai nostri scopi[1].

[*] Questa ricerca è stata realizzata all'interno dei Progetti HUM2007-60410/FILO, «Textualización y oralidad: del español clásico a nuestros días», finanziato dal Ministerio de Ciencia y Tecnología del Gobierno de España, e P08-HUM-03561, «Conciencia Lingüística y Usos Idiomáticos en la Andalucía de la Era de la Información», finanziato dalla Junta de Andalucía (Proyectos de Excelencia), sviluppati dal gruppo di ricerca EHA (El español hablado en Andalucía) (HUM-134). Ringrazio Margarita Borreguero Zuloaga, Covadonga Fouces González, Claudia Collufio e Antonella d'Angelis per il lavoro di traduzione e revisione del testo.

[1] Cfr. Bambini (2005: 949): «Since spoken language has a fragmented structure, characterized by topicalizations, dislocations, ellipsis, etc., the traditional (formal)

1.1 L'approccio del *Groupe Aixois de Recherche en Syntaxe* (*GARS*)

Per evitare quest'orientamento scritturale, intrinseco all'uso delle unità sintattiche di analisi grammaticale (tradizionale o meno, ma basato preferenzialmente su costruzioni prototipicamente scritte) e tipico dei primi approcci alla sintassi della lingua parlata – che mettevano in rilievo la minore complessità di questa in confronto con la sintassi più elaborata delle interazioni linguistiche proprie delle situazioni di distanza comunicativa – emerge, nel seno del *Groupe Aixois de Recherche en Syntaxe* (*GARS*) diretto in un primo momento da Claire Blanche-Benveniste, la proposta di spiegare l'andamento sintattico colloquiale con l'aiuto delle figure della sintassi.
Grazie a queste figure, che rendono possibile l'analisi delle tracce del processo di produzione – che l'edizione dei testi scritti cancella, ma che vengono conservate dalla registrazione e trascrizione (a fini linguistici) delle produzioni orali –, si possono scoprire schemi supraindividuali di costruzione del discorso non pianificato, cioè meccanismi generali per far progredire il discorso e commentare, come di solito avviene, lo sviluppo della composizione.
Questi schemi non sono esclusivi di una lingua specifica, si tratta piuttosto di schemi universali di verbalizzazione, attivati in situazioni di comunicazione immediata (cfr. Koch and Oesterreicher, 1985; 1990/2007; 2001). Alcuni di essi sono più sensibili a determinati parametri situazionali la cui incidenza diminuisce man mano che ci allontaniamo dalla massima immediatezza comunicativa, mentre altri sono presenti anche in discorsi fortemente pianificati (ma pianificati consapevolemente per essere realizzati oralmente e recepiti auditivamente).
Tra le tracce di produzione discorsiva che fanno emergere le figure di sintassi è fondamentale la frammentazione del *continuum* delle unità linguistiche, tipica dello sviluppo lineare del discorso immediato, dovuta all'intersezione dell'asse paradigmatico nel sintagmatico.
Per identificare le aree di intersezione paradigmatico-sintagmatica, Blanche-Benveniste e la sua scuola eseguono la cosiddetta "analisi *en grilles*". Essa consiste in una rappresentazione grafica del materiale

---

syntactic terms do not properly suit it». Lo stesso fatto è denunciato da Cresti (2000: 20-1).

trascritto che permette di visualizzare i fenomeni caratteristici del discorso in costruzione attraverso l'uso dell'asse verticale. Le sequenze nelle quali il paradigmatico si manifesta nel sintagmatico o nelle quali la produzione del discorso include movimenti di andata e ritorno lungo l'asse sintagmatico si trascrivono nella disposizione verticale che si associa iconicamente con l'asse della selezione. Si tratta, in definitiva, di un modello di analisi sintattica basato sulla scrittura a colonna di elementi che corrispondono all'asse paradigmatico, anche se nel parlato appaiono ovviamente in successione. In questo modo, sequenze discorsive come (1a) e (2a) si rappresentano in griglie, per la loro analisi posteriore, come vediamo in (1b) e (2b)[2]:

(1a)
ROS: ya tenéis todo?
MIG: no //
PAT: no // nos faltan / doscientas mil pelas / <pelas> //
MIG: <doscientas> / setenta mil pelas // [efamcv01, 40-44]

(1b)

no						
no	*nos*	*faltan*	*doscientas*		*mil*	*pelas*[3]
			doscientas	setenta	mil	pelas

(2a)
MAX: questo è il deserto //
ELA: <questo è il &fa> [/] famoso deserto // [ifamcv01, 151-152]

(2b)

questo	è	il		deserto
questo	*è*	*il*	*fa-*	
			famoso	*deserto*[4]

[2] Negli esempi identificati con la lettera *a* si rispettano le convenzioni di trascrizioni del C-ORAL ROM. Al riguardo, si veda, oltre Cresti and Moneglia (2005), Nicolás (2008). Nell'analisi in griglie, quando interviene più di un parlante, le parole del primo appaiono in tondo, quelle del secondo in corsivo, quelle del terzo in maiuscoletto, quelle del quarto in grassetto ecc.
[3] Figura della scala (cfr. 3.5.).
[4] Figura della scala (cfr. 3.5.).

1.2 Obiettivi

All'interno di questa linea di ricerca, in questo lavoro tenterò di eseguire, con l'aiuto delle figure di sintassi elencate, l'analisi del materiale in spagnolo e italiano parlato raccolto nel C-ORAL ROM. *Integrated Reference Corpora for Spoken Romance Languages* (Cresti and Moneglia, 2005).
Questo corpus, che è già stato impiegato per altri studi (cfr. per es. Nicolás, 2008), risulta particolarmente interessante per dimostrare il modo in cui parlanti di lingue diverse fanno ricorso agli stessi modelli di costruzione del discorso improvvisato, senza previa pianificazione, in condizioni di massima immediatezza comunicativa.
Grazie all'ampia varietà di situazioni comunicative di interazione orale, ma con diversi profili concettuali (dalle conversazioni private ai materiali provenienti dai *media*), è possibile rintracciare in questo corpus, come abbiamo anticipato, la diminuzione dell'incidenza di determinati tipi di figure di sintassi man mano che ci allontaniamo dalla massima immediatezza comunicativa, mentre altri tipi sono presenti anche nei discorsi molto elaborati (ma pianificati in vista della loro realizzazione orale e di una ricezione auditiva).
Nel presente lavoro documento in questo corpus romanzo le figure di sintassi menzionate e poi completo l'analisi esclusivamente formale della proposta originaria del *GARS* con una classifica funzionale delle diverse figure di sintassi individuate, allo scopo di capire meglio gli obiettivi comunicativi ai quali corrispondono i diversi schemi universali di costruzione del discorso immediato.
In ricerche future sarà necessario rintracciare la loro presenza in discorsi rappresentativi della distanza comunicativa per poter offrire una spiegazione sulle cause della scomparsa di alcune di queste figure di sintassi e la contemporanea presenza di altre (anche se meno frequenti) nei discorsi che non avvengono in condizioni di massima immediatezza comunicativa.

2 Figure di sintassi: classifica formale

L'aspetto formale che le figure di sintassi presentano nel riflettersi dell'asse paradigmatico sull'asse sintagmatico non è completamente uniforme. In uno dei suoi primi lavori (pubblicato in spagnolo), Blanche-Benveniste parla di quattro tipi di figure: *figura di simmetria*,

figura di enumerazione, *figura della parentesi* e *figura del fattore comune*. A queste mi è sembrato necessario aggiungerne un'altra, chiamandola *figura della scala* (cfr. López Serena, 2007b; 2008; 2009). Vediamo brevemente le loro definizioni e alcuni esempi rappresentativi.

2.1 Figura di simmetria

La figura di simmetria è caratterizzata dalla ripetizione di materiali sintattici e lessicali (simile alle figure retoriche come l'anafora, l'epifora, la geminazione, l'epanadiplosi o i parallelismi sintattici[5], in disposizione identica o a forma di chiasmo[6]), che possono provenire sia da un unico emittente come in (4), o da due diversi come in (3):

(3)
GUI: ¡bah! / no seáis también... si no pasa nada // o sea te [/] molesta el aparato ?
[...]

[5] Due esempi consecutivi di parallelismo sintattico li troviamo nel seguente passo – che apparirà sotto come esempio (14) –. Esempio tratto dalla parte italiana del corpus C-ORAL ROM (per esempi spagnoli, cfr. López Serena (2007b):
MAX: questa è un anno prima / e questa un anno dopo //
LIA: dici?
MAX: questa è cinquantanove / e questa è sessanta //
LIA: dici? [ifamcv01, 239-242]

	questa	È	un	anno	prima
E	questa	È	un	anno	dopo

	Questa	è	cinquantanove
E	Questa	è	sessanta

[6] Una simmetria inversa o chiasmo, anche se non totale, la troviamo nel seguente frammento, riprodotto sotto come esempio (10):
LIA: sì / la faccia è giovane / però / <vedi> //
ELA: [<]<sì> / xxx //
LIA: la faccia è giovanile / vedi // vedi / la faccia è giovanile // eh // [ifamcv01, 187-191]

Sì	la	faccia	è	giovane	peró	vedi				
Sì	la	faccia	è	giovanile		vedi				
						vedi	la	faccia	è	giovanile

MIG: te molesta el aparato ? [efamcv01, 8-14]

te	molesta	el	aparato
te	*molesta*	*el*	*aparato*

(4)
LIA: [<]<quando stavano> insieme // che giela portò / a lei / no? per farsi perdonare / che+ che si doveva fa' perdonare / non l'ho mai voluto sapere //
ELA: eh / birbone //
LIA: capisci? no no / non l'ho mai voluto sapere // [ifamcv01, 18-25]

che	si	doveva	fa'	perdonare	non	l'ho	mai	voluto	sapere
					no				
					no				
					non	l'ho	mai	voluto	sapere

2.2 Figura di enumerazione

Come nelle figure retoriche di accumulazione, si tratta di casi di accatastamento paradigmatico nell'asse sintagmatico dovuti alla giustapposizione o coordinazione di almeno due elementi (parole, frammenti di parole o sintagmi completi) non identici che occupano una stessa posizione sintattica. Per quanto riguarda il suo aspetto formale, questa figura può essere il risultato di una semplice enumerazione semantica come in (5), oppure di una falsa partenza come in (6), o ancora della successione di forme sintattiche incomplete che si completano lungo lo sviluppo dell'enumerazione (7). In quest'ultimo caso è meglio parlare di sequenze nella costruzione dovute a una figura di enumerazione (dato che esistono anche le sequenze in costruzione dovute a figure di simmetria (cfr. 8) o di scala).

(5)
MAX: questa sarà di' cinquantanove / e avrà avuto cinquasett' anni // avea la tua età // [ifamcv01, 95-96]

questa	sarà	di'	cinquantanove	e	avrà	avuto	cinquasett'	anni	
						avea	la	tua	età

(6)[7]
*ELA: [129] questa / tu sei a Parigi //
[...]
*ELA: [131] sei con la torre Eiffel / dietro //
*LIA: [132] sì?
*ELA: [133] sì // [134] guarda //
*LIA: [135] sie // [146] sì / sì / sì // [137] siamo [/] so' a Parigi / costì //

questa	tu	**SEI**	**a**	**Parigi**			
		SEI	**con**	**la**	**torre**	**Eiffel**	**dietro**
sì?							
sí	guarda	*sie-*					
sì							
sì							
sì		*siamo*					
		so'	***a***	***Parigi***	*costì*[8]		

(7)
LIA: no no no no no // queste / son persone che non conosco // questa / gliela devono aver data / alla mi' mamma //
ELA: persone che <conosceva / la tua mamma>?
LIA: [<]<qui non c'è &ness>+ può darsi / che gliel' abbino data / perché qui non c'è nulla // qui 'un c'è mamma / 'un c'è nessuno // di persone che conosco // [ifamcv01, 252-259]

									queste	sono	persone	che	no n	conosco			
questa		gliela	devono aver	data	alla		mi'	mama									
											persone	*che*		*conosceva*	*la*	*tua*	*mamma*
						qui	non	c'è	ness-								
può	darsi	gliel'	abbino	data	perché[9]	qui	non	c'è	nulla								
						qui	'un	c'è	mamma								
							'un	c'è	nessuno	di	persone	che		conosco			

3.3 Figura della parentesi

Chiamiamo figura della parentesi la configurazione sintattica dei sintagmi che, in modo parentetico, si situano all'interno di un altro

[7] Sottolineo nella trascrizione in griglie la figura di enumerazione con falsa partenza e metto in grassetto la figura dell'enumerazione semantica, che contiene anche un'anafora, cioè una figura di simmetria (in maiuscolo). C'è anche un'altra figura di simmetria nella prima colonna, che non c'è bisogno di mettere in rilievo, dato che è individuabile a prima vista.
[8] Figura della scala con enumerazione al posto di simmetria.
[9] Figura del fattore comune.

sintagma, interrompendolo a modo di commento, come si osserva in (8) e (9).

(8)
MAR: pues igualito ese tanatorio / que los que hay aquí en Madrid // pues era el tanatorio que tienen para los [/] todos los pueblos de [/] de + [cómo se llama eso? del <ése / norte> //
LUI: [<]<de> la sierra //
MAR: de la sierra norte // [...] [efamcv02, 119-124]

pues	era	el	tanatorio	que	tienen	para		los						
							todos	los	pueblos	de[10]				
										de	cómo	se	llama	eso
										de	ése		norte	
										de	la	sierra		
										de	la	sierra	norte[11]	

(9)
ELA: [<]<fai vedere [/] fai vedere> bellino // questo è Erik / e questo / è quello [///] Mark si chiama [ifamcv01, 561-562]

fai	vedere							
fai	vedere	bellino[12]		questo	è	Erik		
			e	questo	è	quello		
						Mark	si	chiama?[13]

3.4 Figura del fattore comune

La figura del fattore comune è la risultante di una determinata formulazione che si trova in relazione sintattica con più di una unità di costruzione. Frequentemente cooccorre con figure di enumerazione nelle quali gli elementi che partecipano all'accatastamento

[10] Sequenza in costruzione con figura di simmetria.
[11] Scala con figura di simmetria dopo una sequenza in construzione con figura di enumerazione.
[12] Figura della scala con figura di simmetria.
[13] Figura della parentesi inserita in una figura di simmetria con parallelismo sintattico e in un'altra di enumerazione a causa dei problemi nella pianificazione improvvisata del discorso.

paradigmatico caratteristico di questa figura sono retti da uno stesso nucleo sintattico, come abbiamo visto in (7).

3.5 Figura della scala

Come in una anadiplosi, in questa figura la produzione graduale del discorso avviene in modo tale che le nuove sequenze successive si agganciano alle sequenze anteriori mediante la ripetizione di uno o più elementi (sia unità lessicali che posizioni sintattiche) a partire dai quali prosegue lo sviluppo dell'enunciato. Nella sua disposizione grafica in griglie, questo tipo di costruzione assomiglia al disegno di una scala, donde il termine coniato. Le figure della scala occorrono di solito contemporaneamente alla figura di simmetria (dato che normalmente emergono dalla ripetizione dell'elemento che serve d'appoggio per la prosecuzione del discorso); tuttavia in alcuni casi possono presentarsi in combinazione con una figura di enumerazione che provochi l'accatastamento paradigmatico di unità non identiche nel punto a partire dal quale continua il discorso. La figura della scala differisce dalla figura di simmetria per due motivi fondamentali. In primo luogo, nelle figure di simmetria si ripetono sequenze complete, di estensione e natura variabile tra la lessia e l'enunciato, mentre nella scala sono frequenti le reiterazioni di elementi inconclusi in una prima enunciazione (*pe- pero*) e di unità funzionali (congiunzioni, preposizioni ecc.). In secondo luogo, nelle figure di simmetria la ripetizione di elementi *non* costituisce un "gradino" nella progressione discorsiva, un meccanismo di tenuta del turno durante la pianificazione dell'intervento, come nel caso delle scale, in cui gli elementi che si ripetono costituiscono punti d'appoggio a partire dai quali si continua la costruzione discorsiva. Inoltre, nelle scale, nella maggior parte dei casi gli elementi che si ripetono presentano due condizioni: da una parte, sono abitualmente particelle funzionali, cioè preposizioni o congiunzioni, o anche pronomi e quasi mai unità con maggiore contenuto lessicale; dall'altra, normalmente si tratta degli elementi che iniziano un nuovo sintagma. Sembra quindi che questo fenomeno sia strettamente collegato alla pianificazione di nuove unità sintattiche precisamente nel loro inizio. Si tratta inoltre di una figura intrinseca ai discorsi che si realizzano oralmente (ossia ai discorsi del *medium* parlato), anche se la loro frequenza è superiore in quelli che non sono stati pianificati previamente (e cioè in quelli che sono

concezionalmente parlati) (si vedano di nuovo gli esempi 1, 6 e 8 sopra).

4 Figure di sintassi: aspetti funzionali[14]

Gli effetti che risultano, secondo il gruppo *GARS*, dall'accatastamento paradigmatico sono diversi: di coordinazione, di precisione o di iperonimia, di reinizio, di ripetizione, di ricerca di parole ecc.
Partendo dalla revisione e rielaborazione delle idee di Tannen (1987a; 1987b; 1989/1992) e Bazzanella (1993; 1994; 1996), tra gli altri autori, sulla forma e le funzioni della ripetizione nella conversazione, mi sembra necessario distinguere tra le seguenti funzioni:
- di enfasi
- di valutazione
- di strutturazione informativa
- semantica
- formulativa

4.1 Figure di sintassi con valore enfatico

Si tratta sia di autoripetizioni sia di alloripetizioni – anche se normalmente sono dovute a un unico interlocutore – di una serie di materiali lessicali e/o sintattici, il cui principale effetto di senso è l'intensificazione (per motivi semantici, stilistici, argomentativi, interattivi...) del contenuto ripetuto. Nei discorsi dialogici hanno pure funzioni interazionali come la presa o il mantenimento del turno di parola, senza perdere però l'effetto intensificatore di base[15]. In altre parole l'intensificazione può essere dovuta a motivi interazionali.

[14] L'esposizione di queste riflessioni funzionali riprende quanto detto in López Serena (2007b: § 4.2.3.1.1.). Tuttavia questa nuova presentazione rivede e riformula alcune delle posizioni presentate in quel lavoro.

[15] Nei casi di autoripetizione prodotti dalla sovrapposizioni di turni di parola, dove le figure di simmetria rispondono al desiderio di mantenere o prendere il turno, la funzione che sembra prevalere (addirittura su quella di enfasi) è quella che Tannen denomina di interazione e Bazzanella conversazionale. Tuttavia bisognerebbe considerare l'effetto di questo tipo di autoripetizioni negli ambiti di produzione (permettono sia di pianificare sia di continuare il discorso) e di comprensione del discorso (avvertono l'interlocutore della rilevanza del contenuto ripetuto) — plurifunzionalità che pertanto rende poco operativa per i nostri fini, come si è detto, la classifica di Tannen —. In realtà è impossibile evitare la plurifunzionalità o

Il prototipo formale è la figura di simmetria (esempi 1[16], 3, 4, 7, 10).

(10)
LIA: sì / la faccia è giovane / però / <vedi> //
ELA: [<]<sì> / xxx //
LIA: la faccia è giovanile / vedi // vedi / la faccia è giovanile // eh // [ifamcv01, 187-191]

sì	la	faccia	è	giovane	peró	vedi				
sì	la	faccia	è	giovanile		vedi				
						vedi	la	faccia	è	giovani

4.2 Figure di sintassi valutative

Le figure di sintassi valutative sono alloripetizioni o autoripetizioni polifoniche che costituiscono reazioni a quanto detto dall'interlocutore (o enunciatario[17], se si tratta di una autoripetizione), siano corroborazioni (11) oppure refutazioni o messe in questione — o

multifunzionalità delle ripetizioni, come segnala Bazzanella, non solo riguardo alle loro diverse macrofunzioni ma anche riguardo alle loro microfunzioni (Bazzanella, 1993: 288), dato che, come indica la stessa autrice in un altro passo, la ripetizione è un complesso fenomeno multidimensionale, la cui polifunzionalità e gradualità rende enormemente difficile il compito di classificare le loro diverse modalità (Bazzanella, 1996: viii). A questo bisogna aggiungere le ulteriori difficoltà dovute alle diverse concezioni dei diversi studiosi che, nelle loro impostazioni teoriche, a volte distinguono radicalmente tra coesione e interazione e a volte considerano ciò impossibile. È questo il caso di Wolfram Bublitz per cui le funzioni di interazione e coesione si fonderebbero in determinate occasioni, per es., in certe ripetizioni favorite dal processo di interazione conversazionale che funzionano come tracce di coesione orientate al ricevente per fargli riconoscere la coerenza discorsiva (Bublitz, 1996: 16). Infatti, se si tiene conto che sia la produzione e la comprensione che la coesione e l'interazione sono procedimenti che contribuiscono alle diverse sfere implicate nella costruzione del discorso, si capisce che a volte è impossibile propendere per l'uno o l'altro aspetto della costruzione testuale.

[16] L'emissione consecutiva dell'avverbio di negazione da parte di MIG e PAT rinforza il valore di verità della risposta negativa che entrambi danno alla domanda di ROS.

[17] Come è noto, il concetto, riferito al soggetto che, indipendentemente da chi siano l'emittente e il locutore, si considera responsabile di un determinato enunciato, procede dalla teoria della polifonia appartenente alla Teoria dell'Argomentazione elaborata inizialmente da Oswald Ducrot e sviluppata successivamente in collaborazione con Jean Claude (cfr. Anscrombre and Ducrot, 1983).

almeno in dubbio — di ciò che viene ripetuto (12)[18]. Il titolo di funzione di valutazione – nuovamente vincolata all'ambito della interazione, ma anche a quello della comprensione – è dovuto al fatto che nonostante la corroborazione, la messa in questione o il commento metalinguistico di quanto è stato detto da un altro interlocutore (o enunciatario) non sono ovviamente equivalenti nella loro totalità. Essi condividono il compito comune di incorniciare il frammento del discorso altrui che si ripete con una nuova sfumatura modale, con la quale il parlante-ricevente partecipa a ciò che viene comunicato dall'altro interlocutore esprimendo a sua volta la sua valutazione.
Il suo prototipo formale è la figura di simmetria, come si vede negli esempi 1[19], 2[20], 6, 15.

[18] Inge Egner (1996: 66) afferma che G. Jefferson (1972) è stata la prima analista della conversazione che ha richiamato l'attenzione sull'alloripetizione interrogativa. In altri corpora dello spagnolo abbiamo trovato anche occorrenze di questa figura che costituiscono commenti di altri tipi, fondamentalmente metalinguistici, preceduti da una ripetizione che esplicitamente assume la forma del discorso indiretto.

(H 38 A 1; cfr. Briz *et al*, 2002)

614 D: bueno ¿de qué?/ ¿por qué no habláis?
615 B: (ERUCTO) [(RISAS) =]
616 A: [(RISAS)]
617 B: = dice *por qué no habláis*/ ROOM (RISAS) así ha sido el PUNTO
618 y seguido↓ nano§

In ogni caso, la retorica tradizionale assegnava già uno status specifico a questo tipo di ripetizioni, ipotizzando che il raddoppiamento si converte in *correctio* quando il secondo elemento ripetuto rettifica quanto dice il primo. L'esempio di Cicerone che seleziona Lausberg (1967: §621) per illustrare la *correctio* «hic tamen vivit: vivit? immo vero etiam in senatum venit» manifesta che, come osserva Ramiro, e d'accordo con le mie conclusioni, «la enmienda se produce por conmutación de la modalidad o actitud del hablante, es decir, de la única manera que permite la propia repetición léxica» (Ramiro, 1995: 85). Concordo anche con quest'autore, quando dice «[m]ientras que las anadiplosis del párrafo [—si intenda prodotte dalle autoripetizioni—] son, en general, intensificadoras, pues repiten la modalidad oracional y el significado referencial, las del dialoguema [—cioè le alloripetizioni—] son frecuentemente rectificadoras, en la medida en que repiten el significado objetivo, pero modifican la modalidad (*correctio*)» (Ramiro, 1995: 88).

[19] Si tratta di una sequenza in costruzione mediante la figura di simmetria nella quale MIG corregge PAT, precisando l'informazione data da quest'ultima .

[20] Si tratta di una sequenza in costruzione mediante la figura di simmetria nella quale ELA precisa le parole di MAX, aggiungendo un aggettivo che modula il contenuto dell'enunciato.

(11)[21]
ELA: dove l'è vicino a+
LIA: sì / sì //
ELA: <a coso>?
LIA: [<]<lì> // [52] lì su [///] a Genova //
MAX: a Recco //
ELA: a Recco // [ifamcv01, 48-54]

(12a)
LIA: ma / li portava abbastanza bene //
MAX: li portava piuttosto male // perché qui è il cinquantanove // [ifamcv01, 100-102]

(12b)

ma	li	portava	abbastanza	bene
	li	portava	piuttosto	male

4.3 Figure di sintassi che contribuiscono a strutturare l'informazione

Per quanto riguarda la struttura informativa, le figure di sintassi possono manifestarsi sotto forma di figure di simmetria che ripetono un determinato tema diverse volte (7, 13), o come parallelismi sintattici che favoriscono la continuazione del filo del discorso sia nell'ambito della produzione sia in quello della ricezione (9, 10). Le une e gli altri si distinguono difficilmente nell'occorrenza di altri tipi di figure, come quelle di simmetria con funzione valutativa o quelle di enumerazione semantica (che vedremo sotto). Per questo motivo risulta preferibile restringere l'uso rigoroso di questa etichetta alla descrizione di figure che svolgono un'altra funzione che sembra essere passata inavvertita agli studiosi del discorso orale (e appare molto più raramente delle altre nei corpora): parliamo della funzione di delimitare il principio e la fine di un determinato tema all'interno di un'interazione nella quale si trattano anche altri argomenti (15).
Il suo prototipo formale è, ancora una volta**,** la figura di simmetria, come si osserva nei seguenti esempi (13, 14, 15):

[21] Si osservi che le figure di simmetria per alloripetizioni corroborativa come questa che si può osservare nell'esempio (12) sono molto vicine a quelle che hanno funzione enfatica, dato che la corroborazione di ciò che viene detto da un altro intensifica il suo valore di verità.

(13a)
LIA: [<]<io dissi / no no no no> no no no / dissi // le cose sono tua / e <le cose sono> [/] e tua rimangano // [ifamcv01, 29-30]

(13b)

io	dissi	no				
		no				
		no				
		no				
		no				
		no				
		no				
	dissi	le	cose	sono	tua	
	e	le	cose	sono		
				e	tua	rimangano

(14a)[22]
MAX: questa è un anno prima / e questa un anno dopo //
LIA: dici?
MAX: questa è cinquantanove / e questa è sessanta //
LIA: dici? [ifamcv01, 239-242]

(14b)

	questa	è	un	anno	prima
e	questa	è	un	anno	dopo

(14c)

	questa	è	cinquantanove
e	questa	è	sessanta

(15)[23]
IVN: eh / uno di loro / déan venire //
GNA: sì / Giampaolo verrà //
SRE: Giampaolo / m'ha detto che alle nove veniva //
LEO: ha parlato con te / <Giampaolo> ?
SRE: [<] <mh> //
[...]
SRE: porta / quella porta / mamma mia //
[aquí hablan sobre una puerta rota que hay en la sala de ensayos (sala prove)]

[22] Come abbiamo avvertito, la funzione di parallelismo in questo esempio è identica a quella delle figure di enumerazione semantica.
[23] Come si vede, SRE usa praticamente le stesse parole al principio della conversazione, prima di iniziare il tema sulla porta, e alla fine, una volta che ha concluso il tema e ritorna alla questione dell'appuntamento con Gianpaolo.

[...]
LEO: [<] <su è aperto>? <ce l' ha / Giampaolo / le chiavi> ?
IVN: [<] <sì / è aperto> //
SRE: comunque / è aperto // io / con <Giampaolo / son rimasto / che> ci si &sen [/] che [/] che [/] ci si vedeva qui alle nove // [ifamcv02, 1-4, 18, 59-62]

4.4 Figure di sintassi semantiche

Le figure di enumerazione rispondono di solito a tre motivi fondamentali: motivi semantici, la ricerca di una parola o espressione, il desiderio di riprendere in un modo diverso la costruzione di una sequenza determinata. In questo modo, la funzione semantica che assolve questa figura non è altra che quella ovvia di enumerare (16). Anche la figura del fattore comune, nella quale, come abbiamo detto, gli elementi che partecipano all'accostamento paradigmatico caratteristico di questa figura sono retti da uno stesso nucleo sintattico, risponde a motivi sintattico-semantici (7, 17). Dunque i seguenti esempi sono allo stesso tempo di figura di enumerazione e figura del fattore comune, perché risulta difficile trovare casi nei quali entrambe non coincidano.
Il prototipo formale delle figure di sintassi semantica sono la figura di enumerazione e la figura del fattore comune (esempi 5, 6, 7, 11, 14, 16, 17).

(16)[24]
LIA: [<] <ah no / Baratti> è qua // già / tu ha' ragione // qualcosa di' genere // che l'ho addosso / questo [/] questo coso hhh // non c'ha la &fo [///] o no &nun [/] 'un mi ricordo / indo' ho messo le fotografie / delle Dolomiti / e della Germania // [...] [ifamcv01, 59-64]

	non					c'ha	la	fo-			
o	no										
	nun										
	'un	mi	ricordo	indo'	ho	messo	le	fotografie		delle	Dolomiti
									e	della	Germania

(17)
ELA: eh / questo è Eno // e qui / lui / quanto c'aveva / Massimo?
LIA: i' babbo / <mi' marito>? [efamcv01, 155-158]

[24] In questo caso, la figura di simmetria semantica finale si produce dopo una serie di autoripetizioni causate dal reinizio e da una sequenza in costruzione.

e	qui			lui
		quanto	c'aveva	Massimo?
				i' babbo
				mi marito

4.5 Figure di sintassi formulative

Le figure di sintassi di cui si serve il locutore per la pianificazione in atto del suo enunciato, che possiamo denominare formulative perché funzionano nell'ambito della formulazione del discorso, sono, fondamentalmente, le figure della scala (1, 2, 4, 6, 8, 9, 11[25], 15[26], 18), le figure di simmetria ed enumerazione che creano sequenze in costruzione (1, 2, 10, 19) o che si producono per reinizi (3, 7, 13, 16, 20), e alcune figure della parentesi che non obbediscono a fini stilistici, ma a problemi di formulazione come la ricerca di una parola o di un dato nella memoria (8, 9, 21).
Non c'è un prototipo formale ben definito per la realizzazione di questa figura, che è legata a una delle caratteristiche più spiccate dell'immediatezza comunicativa: la realizzazione improvvisata del discorso. Può apparire in figure che normalmente sono della scala, di simmetria e di enumerazione che creano sequenze in costruzione, ed è anche presente in alcune figure della parentesi.

(18)
ELA: qui tu se' [/] qui tu se' piccino // eh / <Massimo>? [ifamcv01, 109-110]

qui	tu	se'			
qui	tu	se'	piccino	Eh	Massimo?

(19)
*ELA: [67] mi ricordo che / tu l'ha [///] me l'ha fatte vedere anche a me //

mi	ricordo	che	tu		l'	ha					
				me	l'	ha	fatte	vedere	anche	a	me

(20)

[25] Mi riferisco a: LIA: [<]<lì> // [52] lì su [...].
[26] SRE: comunque / è aperto // io / con <Giampaolo / son rimasto / che> ci si &sen [/] che [/] che [/] ci si vedeva qui alle nove.

MAX: accidenti // e questo è il babbo / quando stavano +
ELA: <eh / a Genova> //
LIA: [<] <quando stavano> insieme // che giela portò / a lei / no? [...] [ifamcv01, 16-20][27]

e	questo	è	il	babbo	quando	stavano	eh						
							a	Genova					
					quando	stavano	insieme	che	gliela	portò	a	lei	no?

(21a)
LIA: eh / sie / sugli scogli di [/] de [/] de+
ELA: io ‘un lo so / perché ‘un lo so / <come si chiamano i posti> //
LIA: [<] <xxx Barà [/] a> Baratti // ’un lo so / dove si chiamavano / dove andavano loro / sempre+ [ifamcv01, 44-47]

(21b)

eh	sie	sugli	scogli	di													
				de													
				de	io	‘un	lo	so	perchè	‘un	lo	so	come	si	chiamano	i	posti
					Barà												
				a	Baratti	‘un	lo	so					dove	si	chiamavano		
													dove		andavano	loro	sempre

5 Conclusioni

La classificazione delle funzioni delle figure di sintassi nel modo qui proposto parte dalla frequenza con la quale le figure di simmetria sono associate a valori di enfasi, valutazione e strutturazione informativa, quelle di enumerazione a compiti semantici e di pianificazione del discorso (funzione formulativa) e quelle della scala alla costruzione del discorso improvvisata (ancora una volta, funzione formulativa). Tuttavia non si tratta né di funzioni che corrispondono unicamente a queste figure, né di funzioni che le suddette figure svolgono esclusivamente, cioè a dire che ogni figura ricopre unicamente una funzione per volta. Come succede anche con i segnali discorsivi, (unità con compito simile a quello delle figure della sintassi), molte delle figure di simmetria, di enumerazione, di parentesi o scala si possono attuare simultaneamente nell’ambito dell’espressione enfatica e in quello della struttura informativa o la pianificazione. Questo è possibile perché non si tratta di funzioni equivalenti che appartengono a una stessa categoria paradigmatica e che si escludono perciò

[27] L’intervento di LIA iniziava nell’esempio (5).

mutuamente, piuttosto si tratta di funzioni appartenenti a diversi livelli di analisi del discorso: l'espressione della soggettività, l'organizzazione dei turni di parola, la formulazione del discorso, il contenuto semantico, la struttura dell'informazione.
Nelle sue proposte in proposito, il *GARS* ha escluso esplicitamente dall'analisi le considerazioni sulle intenzioni dei parlanti, ritenendo che si tratta di una questione ardua, di difficile approccio. Il GARS si è anche rifiutato di ammettere la differenza tra configurazioni sintattiche prodotte da un unico parlante o da diversi parlanti, nella misura in cui ritiene che un dialogo costituisca un unico discorso, risultante dalla costruzione cooperativa tra diversi parlanti, i cui contributi individuali scompaiono a favore di un prodotto unitario. Tuttavia se, come nel nostro caso, si desidera mostrare la redditività che nello studio della variazione *concezionale* parlato/scritto possa avere la possibilità di rintracciare in un corpus come il C-ORAL ROM la diversa incidenza che i tipi diversi di figure di sintassi manifestano in discorsi più o meno prossimi al polo di massima immediatezza comunicativa, risulta imprescindibile tenere conto di entrambi gli aspetti. Anzi, per interpretare le diverse funzioni che svolgono i diversi tipi di ripetizione non è neanche imprescindibile esprimere giudizi o congetture sulle intenzioni nel produrre o meno le diverse figure sintattiche che vengono esemplificate in questo studio. Si tratta piuttosto di rivolgere l'attenzione all'effetto discorsivo o pragmatico di queste figure, che in grande misura è in correlazione con il fatto che esse procedano dall'emissione di uno o più parlanti e che siano intersoggettivamente interpretabili.
Come è accettato, almeno da Halliday and Hasan (1976), nei vari studi sull'organizzazione suprafrasale del discorso, la ripetizione di elementi lessicali e/o sintattici (che in un modo o in un altro conformano la base delle figure di simmetria, enumerazione, fattore comune e scala) costituisce un procedimento di coesione discorsiva[28]. Ma simultaneamente alla realizzazione del compito coesivo, le figure di sintassi assolvono altre numerose funzioni di diverso tipo. In lavori precedenti abbiamo mostrato la redditività delle considerazioni

[28] D'accordo con Bublitz (1996: 19 e *n.* 8), anche se Brown (1977: 113 *s.*), come Hasan and Halliday, segnalò prontamente la funzione coesiva della ripetizione, fu Johnstone (1991) la prima ad insistire realmente sulla funzione di strutturazione del discorso dei *pattern* di ripetizione.

formali e funzionali delle figure qui trattate per lo studio della mimesi dell'oralità nella letteratura (cfr. López Serena, 2007b) o come prova della sempre maggiore oralizzazione dei media audiovisivi (López Serena, 2008; 2009). In questo nuovo lavoro, il mio obiettivo è stato offrire una breve esemplificazione della loro possibile utilità per la ricerca sulla immediatezza comunicativa in lingue come l'italiano, con l'augurio di suscitare l'interesse per il loro studio nell'italiano parlato, dove, per quanto riguarda le mie conoscenze, non sono ancora state prese in considerazione, a differenza di altri aspetti proposti da Blanche-Benveniste e la sua scuola.

Riferimenti bibliografici

Anscombre, J. C. and O. Ducrot (1983) *L'argumentation dans la langue*. Bruxelles: Mardaga.

Bambini, V. (2005) Review of Corpus di italiano parlato: Vol. 1: Introduzione, Vol. 2: Corpora, CD-Rom Emanuela Cresti, Accademia della Crusca, Firenze, 2000. *Journal of Pragmatics* 37, pp. 949-53.

Bazzanella, C. (1993) Dialogic repetition, in H. Löffler, Ch. Grolimund and M. Gyger (eds) *Dialoganalyse IV. Referate der 4. Arbeitstagung Basel 1992*, vol. 1, pp. 285-94.

– (1994) *Le facce del parlare: un approccio pragmatico all'italiano parlato*. Scandicci (Firenze): La Nuova Italia Editrice.

– (ed.) (1996) *Repetition in Dialogue*. Tübingen: Max Niemeyer Verlag.

Blanche-Benveniste, C. (1985) Las regularidades configurativas en el discurso del francés hablado. Consideraciones lingüísticas y sociolingüísticas, in F. Rodríguez Izquierdo (ed.) *Sociolingüísitca andaluza*, 3. *El discurso sociolingüístico*, Sevilla: Servicio de Publicaciones de la Universidad de Sevilla, pp. 19-30.

Briz, A. *et al* (2002) *Corpus de conversaciones coloquiales*. Madrid: Arco Libros (Anejo I, *Oralia*).

Brown, G. (1977) *Listening to Spoken English*. London: Longman.

Bublitz, W. (1996) *I bought a freezer – you've bought a freezer – they've bought a freezer*. Repetition as a Text-building Device, in C. Bazzanella (ed.), pp. 16-28.

Cresti, E. (2000) *Corpus di italiano parlato. I. Introduzione. II. Campioni*. Firenze: Accademia della Crusca.

Cresti, E. and M. Moneglia (eds) (2005) *C-ORAL-ROM. Integrated Reference Corpora for Spoken Romance Languages*. Amsterdam: John Benjamins.
Egner, I. (1996) Other-repetition in Question form: Evidence from a West-African Language, in C. Bazzanella (ed.), pp. 66-77.
Halliday, M.A.K. and R. Hasan (1976) *Cohesion in English.* London: Longman.
Jefferson, G. (1972) Side sequences, in D. Sudnow (ed.) *Studies in Social Interaction*, New York: Free Press, pp. 294-338.
Johnstone, B. (1991) *Repetition in Discourse*. Norwood, N.J.: Ablex, vol. 1.
Koch, P. and W. Oesterreicher (1985) 'Sprache der Nähe – Sprache der Distanz. Mündlichkeit und Schriftlichkeit im Spannungsfeld von Sprachtheorie und Sprachgeschichte'. *Romanistisches Jahrbuch* 36, pp. 15-43.
Koch, P. and W. Oesterreicher (2001) Gesprochene Sprache und Geschriebene Sprache/Langage parlé et langage écrit, in G. Holtus, M. Metzeltin & Ch. Schmitt (eds) *Lexikon der Romanistischen Linguistik*, vol. ½, Tübingen: Niemeyer, pp. 584-628.
– (1990/2007) *Lengua hablada en la Romania: español, francés, italiano*. Madrid: Gredos [trad. esp. (1990) A. López Serena de, *Gesprochene Sprache in der Romania: Französisch, Italienisch, Spanisch*. Tübingen: Niemeyer].
Lausberg, H. (1967) *Manual de retórica literaria.Fundamentos de una ciencia de la literatura* [trad. di José Pérez Riesco. Madrid: Gredos, 3 voll.].
López Serena, A. (2007a) 'El concepto de 'español coloquial': vacilación terminológica e indefinición del objeto de estudio'. *Oralia* 10, pp. 161-91.
– (2007b) *Oralidad y escrituralidad en la narrativa literaria*. Madrid: Gredos.
– (2008) La escritura(liza)ción de la sintaxis oral en la edición de entrevistas periodísticas, in R. Schmidt-Riese, E. Stark & E. Stoll (eds) *Romanische Syntax im Wandel*, Tübingen: Narr, pp. 531-47.
– (2009) Los medios de comunicación audiovisual como corpus para el estudio de la sintaxis coloquial, in M.ª Camacho Taboada, J. J. Rodríguez Toro & J. Santana Marrero (eds) *Estudios de lengua española: descripción, variación y uso*, Madrid-Frankfurt: Iberoamericana-Vervuert, pp. 405-37.

Nicolás Martínez, C. (2008) Primeros resultados del estudio de *nada*, *niente* y *nulla* en C-ORAL-ROM, in C. Nicolás Martínez (ed.) *Ricerche sul Corpus del parlato romanzo C-ORAL-ROM. Studi linguistici e applicazioni didattiche per l'insegnamento di L2*. Firenze: Università, available online from http://electronica.unifi.it/online/coralrom/ (accessed: 9 September 2009), pp. 49-66.

Ramiro Valderrama, M. (1995) *El énfasis en la prosa de Cela. La repetición como procedimiento connotativo*. Valladolid: Servicio de Publicaciones de la Universidad de Valladolid.

Tannen, D. (1987a) 'Repetition in conversation: Toward a poetics of talk'. *Language*, 63, 3, pp. 574-605.

– (1987b) 'Repetition in conversation as spontaneous formulaicity'. *Text* 7, 3, pp. 215-43.

– (1989/1992) *Talking Voices: repetition, dialogue, and imagery in conversational discourse*. Cambridge: Cambridge University Press.

Spontaneous speech intonation in Italian
An experimental analysis with a macrosyntactic view

Philippe Martin
UFRL Université Paris Diderot
philippe.martin@linguist.jussieu.fr

Abstract

Most studies on Italian intonation pertain mainly to read speech, and use for a large part either an autosegmental-metrical framework (Avezani and Vayra, 1994) or, for a small number, a phonosyntactic model (Martin, 2004). In papers pertaining to the last category, there is a clear assumption that the sentence prosodic structure is independent but associated to syntax. In this view, prosodic contours located on or around stressed syllables function as phonological markers of this prosodic structure. Experimental studies describe for example a high and rising melodic contour located on the first stressed syllable of a subject NP, and either a rising contour (on the stressed syllable) or a complex contour (falling on the stressed syllable and rising on the last syllable) of the last unit of a SN syntagm.
Whereas this description is essentially compatible with descriptions given in the autosegmental-metrical (AM) framework, the validity of AM theory may be questionable if extended to non prepared (spontaneous) speech. In such cases, a macrosyntactic approach proves to be an effective tool, as it envisions the sentence as a sequence of macrosegments, syntactically well formed in the classical sense, but whose relations of parataxis or dependency with each other are partially determined by the sentence prosodic structure. This paper presents a short example of analysis conducted in this framework, showing the interaction between macrosyntactic and prosodic structures.

1 Introduction

Macrosyntactic analysis (Blanche-Benveniste, 2000) proceeds by segmenting the text in various macrosegments, which can be defined roughly as delimited to the left (for the left boundary) and to the right (for the right boundary), by an absence of dependency being considered strictly from a syntactic point of view (excluding semantic dependencies). One of these macrosegments, called the kernel, is special as it can form a complete well formed sentence both on the

syntactic and prosodic levels. This implies it can be isolated (with a sound editor for example) and perceptively appears as complete. Other tests, implying a change in modality for example, can also be conducted in order to assert the character of kernel of a given macrosegment.
Macrosegments appearing before the kernel are called prefixes, those inside the kernel parentheses, and macrosegments placed after the kernel are either suffixes or postfixes, depending on the presence or absence of an explicit syntactic relation with the kernel that precedes.

2 An example

The example analyzed below was taken from the C-ORAL-ROM corpus (Cresti and Moneglia, 2005), file IFAMDL09, and consists of one female Tuscan speaker SAB talking to a friend about her (painful) attendance to a concert the night before. Transcribed without any punctuation, the text appears as follows:

**SAB: ma io penso non so quanta capienza c'ha comunque io penso dodicimila persone s'era tutte mh guarda non cascava uno spillo sinceramente e poi tra l'altro c'è stata la menata che noi ci s'aveva il biglietto per il parterre praticamente quando siamo entrati hanno aperto i cancelli già in ritardo alle sei e mezzo sicché noi siamo stati due ore lì a aspettare in fila così quando siamo entrati c'hanno detto che praticamente noi non si poteva andare sulle gradinate a sedere ma soltanto in mezzo si poteva stare sicché a me mi girava un po'le palle perché insomma stare ancora a aspettare fino alle nove e poi tutto il concerto in piedi insomma era stressante la cosa e poi in piedi hai visto anche se il palco è un po'rialzato però se ti viene uno davanti alto non vedi nulla specialmente io che non sono ba insomma che son bassa vero sicché nulla io e quest'altra ragazza che era in macchina con me s'è detto sai sicché proviamo a andare nelle gradinate e siamo riuscite a sgamare sicché siamo siamo andate su e nulla ci siamo messe a sedere però logicamente tutti i posti erano prenotati*

2.1 Analysis in macrosegments

Spontaneous speech production is often characterized by the presence of hesitations, repetitions, reformulations and ponctuants normally not

found in read speech (i.e. read from written text). In the short sample, we find two reformulations (***ma iO pEnso … comUnque io pEnso, specialmente io che non sono ba insomma che son bassa***), two hesitations (***mh, siaaaaaaaamo***), one repetition (***siamo siamo***) and one ponctuant (***vero***).

2.2 General characteristics

However, the main characteristics of non prepared speech pertain to the macrosyntactic organization of the text in prefixes, kernel, imbedded (infixes, parenthesis), postfixes and suffixes.

The text of the example has been segmented in macrosegments with the following conventions:

Prefix	*italic*
Kernel	**bold underlined**
Imbedded	underlined
~~*Postfix*~~	~~*italic strikethrough*~~
Suffix:	*italic underlined*

Table 1: Conventions

Stressed vowels are indicated by upper case characters. Macrosegments boundaries are indicated by #.

SAB: ma io penso #* non so quanta capienza c'ha *# comunque io penso # dodicimila persone* **s'era tutte *# mh guarda #* **non cascava uno spillo** *#* sinceramente *# e poi # tra l'altro # c'è stata la menata che noi ci s'aveva il biglietto per il parterre # praticamente # quando siamo entrati#* **hanno aperto i cancelli già in ritardo alle sei e mezzo sicché noi siamo stati due ore lì a aspettare in fila così** *# quando siamo entrati #* **c'hanno detto che** *#* praticamente *#* **noi non si poteva andare sulle gradinate a sedere** *#* *ma soltanto in mezzo si poteva stare # sicché a me mi girava un po'le palle # perché #* insomma *# stare ancora a aspettare fino alle nove e poi tutto il concerto in piedi #* insomma *#* **era stressante la cosa** *# e poi in piedi hai*

visto anche se il palco è un po'rialzato # però se ti viene uno davanti alto # **non vedi nulla** *# ~~specialmente io che non sono ba insomma che son bassa vero~~ # sicché nulla io e quest'altra ragazza che era in macchina con me # s'è detto# sai#* **sicché proviamo a andare nelle gradinate e siamo riuscite a sgamare** *# sicché siamo # siamo andate su # e nulla #* **ci siamo messe a sedere però logicamente tutti i posti erano prenotati** *#*

Reduced to its macrosyntactic kernels, the text becomes:

s'*era tutte*
non cascava uno spillo
hanno aperto i cancelli già in ritardo alle sei e mezzo sicché noi
siamo stati due ore lì a aspettare in fila così
noi non si poteva andare sulle gradinate a sedere
era stressante la cosa sicché proviamo a andare nelle gradinate e siamo riuscite a sgamare
ci siamo messe a sedere però logicamente tutti i posti erano prenotati

2.3 Intonation macrosegments

The recording of the sample was analyzed with the speech analysis software WinPitch, and the resulting fundamental frequency curves segmented so that the melodic contours located on stressed syllables appear in darker color.

2.3.1 Kernels

As expected, low level and rather falling of flat melodic contour appears on the last stressed syllable of every kernel.

2.3.2 Prefixes

Large prefix melodic contour realized either 1) by a rising melodic contour on the last stressed syllable of every prefix or 2) by a flat of slightly falling melodic contour on the stressed syllable and a sharp rising melodic contour on the last syllable. If the last syllable is also stressed, both melodic movements (flat-falling and rising) take place

on the syllable. The second possibility is mostly found in large (i.e. containing many stress groups) prefixes;
As in read speech, the first stress syllable carries a rising contour.
Inside the prefixes, if there is no sub grouping of stress groups, the prosodic events are neutralized and realized as flat melodic contours.

2.3.3 Parentheses

Parentheses can be associated with an independent prosodic structure (as shown Fig. 1) or be prosodically integrated with the overall sentence prosodic structure. In the first case, the final contour is falling, as it would be the case in a declarative kernel, in the second, the final contour would be rising as in prefixes.

2.3.4 Postfixes

Postfixes carry a sequence of flat melodic contours.

2.3.5 Suffixes

Suffixes are associated with a standard declarative (interrogative) prosodic structure. Their terminal contour is thus falling (rising).

3 Detailed analysis

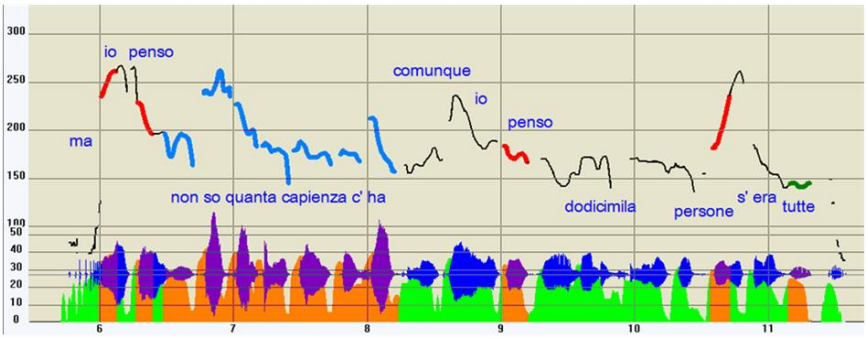


Figure 1: *ma iO pEnso* # <u>non so quanta capiEnza ci hA</u> # *comUnque io pEnso* # *dodicimila persOne* **<u>s'era tUtte</u>** #

Fig. 1 shows a prefix, interrupted by the parenthetic macrosegment non so quanta capiEnza ci hA ended by a falling contour, and followed by the kernel **s'era tUtte**. As in read speech, the first stressed syllable *io* carries a rising melodic contour (Martin, 2004). The macrosegment *ma iO pEnso* is interrupted by the parenthesis non so quanta capiEnza ci hA ended by a terminal falling declarative contour. On the first and second stressed syllable of *ma iO pEnso* and the stressed syllable

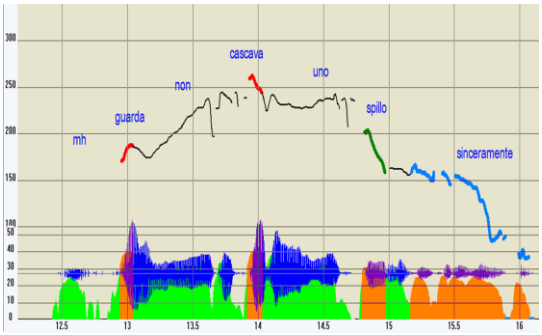


Figure 2: *mh guArda* # **non cascAva uno spIllo** # *sinceramEnte* #

Fig. 2 displays the melodic curve of a prefix followed by a kernel and a suffix. The kernel and the suffix are both ended by a terminal declarative contour on their last stressed syllable.

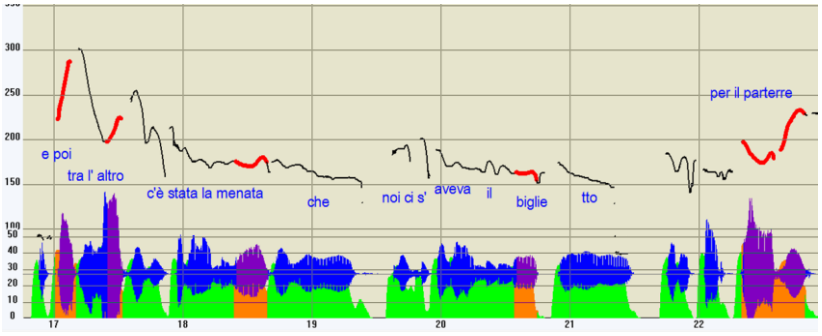


Figure 3: *e pOi # tra l'Altro # c'è stata la menAta che noi ci s' avEva il bigliEtto per il partErre #*

Fig. 3 shows two short prefixes terminated by a rising contour, followed by a long prefix ended by a complex melodic contour, falling on the stressed syllable and rising on the last syllable of *parterre*. This complex melodic contour is often found in long NP in read speech (Martin, 2004).

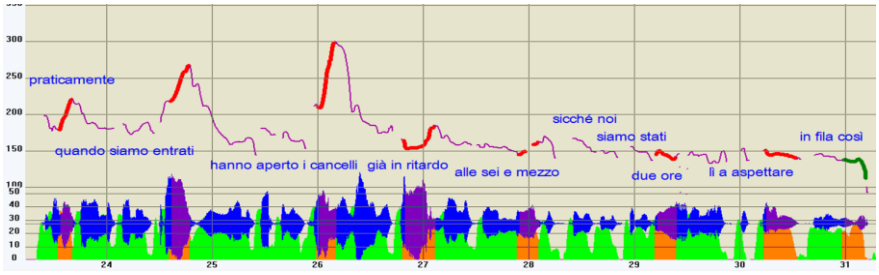


Figure 4: *praticamEnte* # *quando siamo entrAti* # **hanno aperto i cancElli già in ritArdo alle sei e mEzzo sicché noi siamo stati due Ore lì a aspettAre in fila cOsì** #

Fig. 4 presents a sequence of 3 short prefixes, ended by rising contours on their last word stressed syllable, followed by a long kernel. An emphatic stress is found on the stressed syllable of *cancelli*, wich cannot be boundary of a prefix.

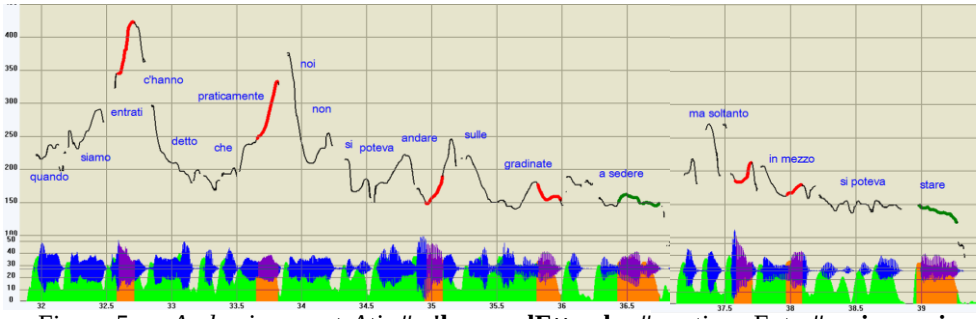


Figure 5: *quAndo siamo entrAti* # **c'hanno dEtto che** # praticamEnte # **noi non si poteva andAre sulle gradinAte a sedEre** # *ma soltAnto in mEzzo* # ~~*si poteva stAre*~~ #

Fig. 5 shows a prefix followed by a kernel interrupted by the parenthesis *praticamente*, the kernel is ended by a terminal declarative contour, low and slightly falling. The kernel is followed by a suffix (*ma soltAnto in mEzzo*) and a postfix (~~*si poteva stAre*~~).

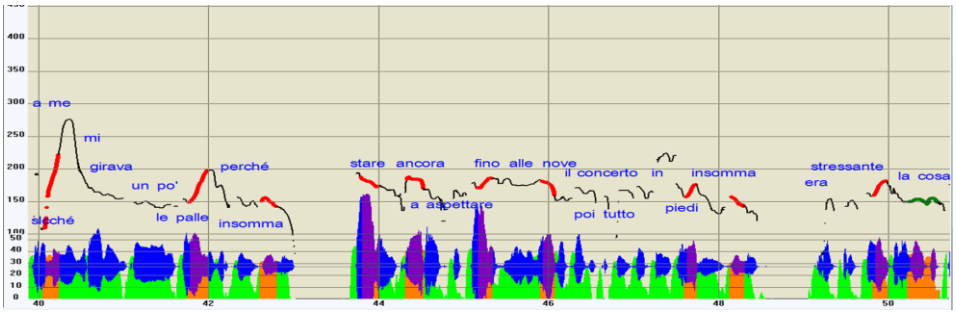


Figure 6: *sicchÈ a me mi girava un po' le pAlle perché # insOmma # stAre ancOra a aspettAre fino alle nOve e poi tutto il concerto in piEdi # insOmma # **era stressAnte la cOsa** #*

Fig. 6 displays a long prefix (9 syllables), interrupted by two parentheses *insOmma* and followed by the kernel **era stressAnte la cOsa**.

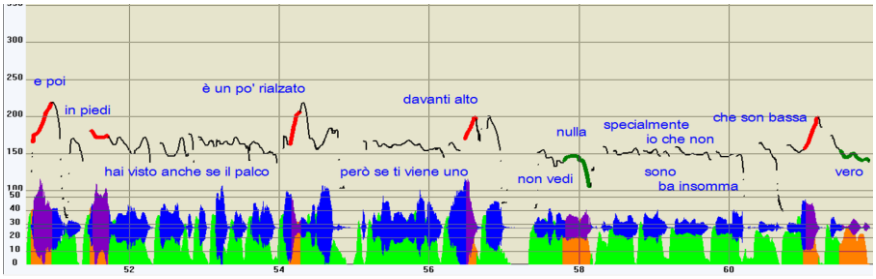


Figure 7: *e pOi # in piEdi hai visto Anche se il pAlco è un po' rialzAto # però se ti viene uno davanti Alto # **non vedi nUlla** # ~~specialmEnte io che non sono ba insOmma che son bAssa vEro #~~*

Fig. 7 shows three prefixes, one kernel (**non vedi nUlla**), followed by a postfix with rather flat melodic contours, at the exception of the

emphatic stress on *bassa*. The end of the postfix is marked by the ponctuants *vero*.

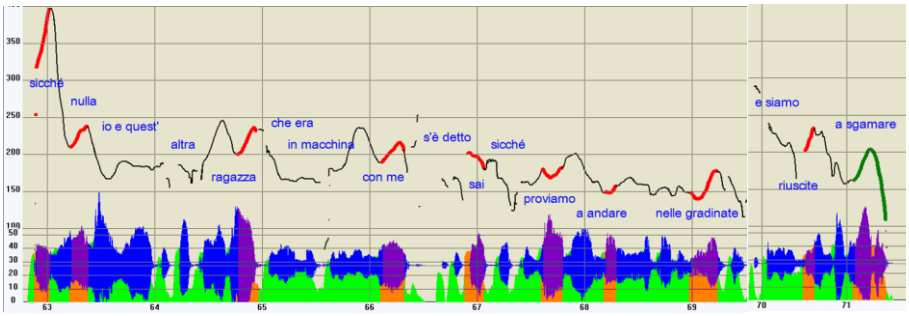


Figure 8: *sicchÈ nUlla # io e quest' altra ragAzza # che era in macchina con mE # s'è dEtto # sAi # **sicchÈ proviAmo a andAre nelle gradinAte e siamo riuscIte a sgamAre** #*

Fig. 8 shows a sequence of 5 prefixes, all ended by a rising contour, whereas the kernel carries a terminal declarative contour, rising on the stressed syllable and falling on the last syllable of the last word *sgamare*.

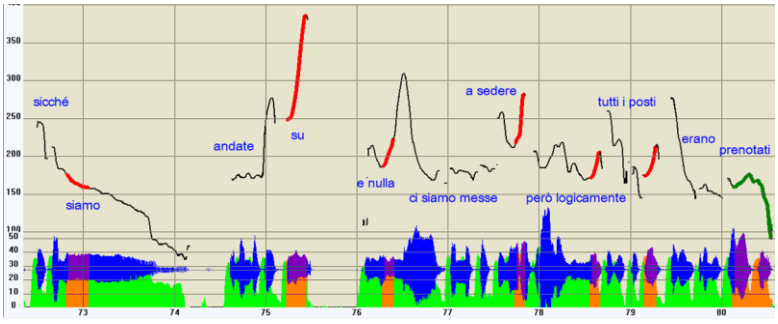


Figure 9: *sicché siAmo # siamo andate sU # e nUlla # **ci siamo messe a sedEre però logicamEnte tutti i pOsti erano prenotAti** #*

In Fig. 9, the sentence does not start with a rising contour on the first syllable, as there is a long filled hesitation on the last syllable of the first occurrence of *siamo*. All prefixes end with a rising contour, and

the kernel ends with a comples final declarative contour, rising on the stressed syllable and falling on the last syllable of *prenotati*.

4 Some conclusions

The following conclusions can be drawn from the analysis of this particular Tuscan speaker:

1. The contour ending the prosodic structure and the kernel shows a flat or moderate melodic rise on the last stressed syllable, and a sharp fall on the last (unstressed syllable);

2. Filled or empty pauses play no significant phonological role. They can be totally absent between two consecutive sentences, or, in the case of filled pause, mark hesitations before reprisal;

3. Rising contours with large melodic excursion occur only at end boundaries of macrosegments (typically with a 70-100 Hz excursion);

4. Inside kernels, stressed syllables are marked with very limited range rising contours (typically in the 20-30 Hz range);

5. Complex contours ending subject Noun Phrase (flat or slightly falling on the stressed syllable and sharply rising on the final syllable) observed in read speech are not frequently used by this speaker, and occur only to mark the end of a relatively long sequence.

Detailed analysis of prosodic data show that the speaker uses efficiently the melodic resources at hand, to mark in a hierarchical way boundaries of macrosegments: the highest level macrosegments in the overall sentence structure are marked by complex contours, then at a lower level by rising contours of large amplitude, whereas inside macrosegments, stress groups are marked by limited range melodic rise, and occasionally at a even lower level by a melodic fall.

References

Avesani, C., and M. Vayra (1994) ToBIt. Un sistema di trascrizione per l'intonazione italiana, in *Atti delle V giornate di studio del Gruppo di Fonetica Sperimentale (GFS)*, 1994, Trento, Italy, pp. 85-98.

Blanche-Benveniste, C. (2000) *Approches de la langue parlée en français*. Paris: Ophrys.

Cresti, E. and M. Moneglia (eds) (2005) *C-ORAL-ROM Integrated Reference Corpora for Spoken Romance Languages.* Amsterdam: Benjamins.

Martin, Ph. (2004) 'L'intonation de la phrase dans les langues romanes : l'exception du français'. *Langue française*, mars 2004, pp. 36-55.

PARTE V

ANTROPOLOGIA E SOCIOLOGIA DELLA COMUNICAZIONE PARLATA

Le formulazioni in contesti istituzionali: un'analisi linguistica e sociologica

Claudio Baraldi
Università di Modena e Reggio Emilia
claudio.baraldi@unimore.it

Abstract

Formulations are types of turns in the interaction which consist in summarising, glossing, or developing the gist of an earlier statement and which, on the one hand, are projected by previous turns, on the other, they project subsequent turns. The analysis of formulations requires attention both for the entire interactional sequence in which they are embedded and for the cultural, presuppositions of the social system in which this interactional sequence is included. The analysis of interactional sequences shows that formulations, and the lexical and syntactic forms in which they are expressed, have different meanings in correlation with different structures of the interaction. The analysis of the cultural presuppositions shows that formulations and the interactional structures which include them, have different meanings in correlation with different forms of expectations which guide participants' actions. These analyses suggest that formulations are actions which can effectively trigger important changes in the interaction, and consequently variations in the social systems in which they occur, variations which could be selected as fundamentals for new cultural presuppositions.

1 Il concetto di formulazione

Il concetto di formulazione (*formulation*) è utilizzato da Heritage (1985) per indicare un tipo di turno che è rilevante in interazioni che si svolgono in contesti istituzionali, come i programmi radiofonici o i tribunali. In queste interazioni, coloro che hanno il ruolo istituzionale di porre domande (ad es., intervistatori radiofonici o avvocati) evitano di posizionarsi anche come ricettori della risposta, per lasciare invece questa posizione agli ascoltatori (ad es., pubblico dei programmi radiofonici, giudici o giurie nei tribunali). Nell'approccio di Heritage, la formulazione rende possibile che un pubblico riceva una risposta.
La formulazione identifica il nocciolo (*gist*) dei turni precedenti e lo riassume, lo commenta oppure lo sviluppa, rendendo esplicito ciò che è in essi implicito, oppure producendo inferenze sui loro presupposti o

sulle loro implicazioni. La formulazione dovrebbe così facilitare la ricezione dell'ascoltatore, aiutandolo a chiarire il senso del turno che ha ascoltato, che può essere troppo complesso, prolisso o dispersivo per essere immediatamente compreso. La formulazione rende anche il contenuto del turno formulato disponibile per una nuova conversazione, poiché proietta una direzione per i turni successivi dell'interlocutore, che può reagirvi e a sua volta contribuire a chiarire il senso di ciò che intende dire, smentire un'inferenza, correggere un riassunto, contestare un commento.

Un'analisi più recente della formulazione non include il pubblico bensì si concentra sull'interlocutore, in interazioni tra psicologi e figli di genitori separati o divorziati (Hutchby, 2007). La formulazione si produce al terzo turno di una tripletta, dopo una domanda dello psicologo e una risposta del bambino (QAF: *question-answer-formulation*). In questa posizione, la formulazione manifesta l'ascolto "attivo" dello psicologo nei confronti del bambino, seguendo un'impostazione tipica della consulenza psicologica. La formulazione è diretta all'interlocutore (il bambino), che dovrebbe cogliervi un "rispecchiamento" (Mearns and Thorne, 1999) della sua posizione discorsiva. Pertanto, la formulazione non è inclusa in una triade (parlante, interlocutore e pubblico) ma in una diade, ed è posizionata al terzo posto in una tripletta. Hutchby (2007) osserva che i bambini tendono a "resistere" alle domande degli psicologi, rispondendo nel modo più breve e semplice possibile: di conseguenza, la formulazione è molto più un'interpretazione che non un rispecchiamento della posizione del bambino. Dal punto di vista della psicologia, che considera l'interpretazione da evitare (Mearns and Thorne, 1999), l'ascolto attivo non è efficace nelle formulazioni analizzate. Dal punto di vista dell'analisi del discorso, i turni dello psicologo non sembrano ottenere risultati brillanti nel proiettare i turni dei bambini: i bambini non si allineano alle domande e quindi non aggiungono nulla di importante a quanto lo psicologo afferma nelle sue formulazioni. L'analisi di Hutchby evidenzia dunque un problema nel rilancio della conversazione attraverso la formulazione, associato ad un modo peculiare di considerare il nocciolo dei turni precedenti.

L'analisi di Hutchby evidenzia tre problemi concettuali: 1) l'identificazione del riferimento primario della formulazione (pubblico o interlocutore); 2) il significato della formulazione nel considerare il

nocciolo dei turni precedenti; 3) il modo in cui la formulazione proietta i turni successivi.

2 L'analisi della formulazione: una proposta

L'identificazione del riferimento primario della formulazione è rilevante per l'analisi. La formulazione può essere interpretata sia come chiarificazione rivolta al pubblico, sia come rilancio rivolto all'interlocutore.

In senso generale, la formulazione è un turno (o un'azione) all'interno di un'interazione. Osservando la formulazione nell'interazione, è subito evidente una tripletta che include: 1) il turno o i turni oggetto di formulazione; 2) la formulazione come turno specifico; 3) la reazione alla formulazione, cioè il suo effetto di *next positioning* (Goodwin and Heritage, 1990). In questa tripletta, il pubblico non è previsto: assiste all'interazione senza potervi partecipare. L'analisi delle conseguenze della formulazione per un pubblico di non partecipanti richiederebbe una verifica della percezione dell'ascoltatore. L'analisi dell'ascolto come tale non riguarda l'interazione in se stessa: per comprendere il significato della formulazione *per* l'interazione, si richiede che la formulazione sia osservata nell'interazione, sia nel rapporto con i turni precedenti, sia nel modo in cui proietta i turni successivi.

Ci sono anche situazioni in cui una formulazione ha destinatari diversi da chi ha preso il turno precedente e che tuttavia partecipano all'interazione complessiva. Ad esempio, la formulazione può essere prodotta in un'interazione di gruppo: oltre ai parlanti che producono il turno formulato e la formulazione, sono presenti altri potenziali parlanti. In questo caso, la formulazione ha un pubblico, al quale è rivolta: tuttavia, questo pubblico è coinvolgibile nell'interazione. Un secondo esempio è l'interazione in cui un partecipante fa da mediatore tra altri due: il mediatore può formulare il contributo di uno dei suoi interlocutori a beneficio dell'altro. In particolare, ci sono studi di casi in cui un'interprete media tra partecipanti che non conoscono l'uno la lingua dell'altro e la formulazione si produce attraverso una traduzione che rende accessibile il nocciolo del turno di un partecipante all'altro, proiettando così l'azione di entrambi (Baraldi and Gavioli, 2007).

L'analisi di queste situazioni può contribuire a chiarire le caratteristiche della formulazione nell'interazione. In particolare,

l'analisi che sarà proposta nella prossima sezione riguarda formulazioni prodotte in classi scolastiche, in gruppi internazionali di bambini e adolescenti e in interazioni mediate da interpreti in servizi sanitari. In questi tre tipi di contesti istituzionali, le formulazioni sono a beneficio di ascoltatori che sono anche potenziali partecipanti: 1) nelle classi scolastiche e nei gruppi internazionali, la formulazione dell'insegnante o dell'educatore va a beneficio dei bambini o degli adolescenti; 2) nei servizi sanitari, la formulazione del mediatore va a beneficio del paziente o del medico.
L'analisi che segue intende evidenziare tre aspetti rilevanti per interpretare la formulazione: 1) la sua collocazione e il suo significato nell'interazione; 2) il suo rapporto con il contesto istituzionale; 3) il cambiamento che determina nell'andamento e nell'orientamento dell'interazione.
L'analisi è condotta su trascrizioni di interazioni videoregistrate (nelle classi scolastiche e nei gruppi internazionali) o audioregistrate (nei servizi sanitari). Tutti i materiali qui utilizzati sono stati già pubblicati in altri saggi e volumi, ma sono utilizzati per la prima volta insieme come materiali utili per un'analisi complessiva della funzione della formulazione. Le formulazioni sono indicate in neretto nelle trascrizioni.

3 L'analisi della formulazione nell'interazione

I dati analizzati evidenziano una varietà di strutture nelle interazioni in cui sono inserite le formulazioni. Dati gli spazi a disposizione, in questa sezione segnaleremo soltanto alcune di queste strutture, le più ricorrenti e importanti.
Una prima struttura di interazione è formalmente analoga a quella rilevata da Hutchby (2007) come QAF: include la formulazione come terzo turno in una tripletta che prevede nei primi due turni una domanda e una risposta. Tuttavia, questa struttura può non avere le caratteristiche segnalate da Hutchby.
La sequenza 1 riproduce una fase di verifica di un'attività di educazione interculturale in una classe del quarto anno di una scuola primaria (Baraldi, 2007). Le domande dell'insegnante, dirette (turni 1, 3, 5, 7) oppure sospese (turno 9), sono seguite da risposte dei bambini e da formulazioni (turni 3, 5, 9) che correggono i contenuti delle risposte, introducendovi variazioni lessicali o sintattiche. La

formulazione sostituisce la correzione esplicita, mitigandone l'effetto valutativo. La mitigazione è segnalata dalle aperture di turno: un segnale di riconoscimento (turno 3: *sì*), una forma verbale condizionale (turno 5: *direi che*) e un eco dell'inizio del turno precedente (turno 9: *che*). Ogni formulazione è seguita da una domanda di verifica come seconda parte del turno: il turno quindi abbina il riassunto della formulazione e lo sviluppo separato dalla formulazione. La formulazione non proietta direttamente il turno successivo, poiché la risposta del bambino è allineata alla domanda di verifica: riprendendo l'analisi di Hutchby, si tratta di una struttura Q-A-FQ (question-answer-formulation+question). Al turno 7, una domanda intermedia dell'insegnante incoraggia il bambino a proseguire e quindi a rendere disponibile il "nocciolo" da formulare, rendendo possibile la formulazione. L'allineamento dei bambini nelle risposte consente all'interazione e alle formulazioni di riprodursi.

S1
Ins. Poi abbiamo letto alcune storie. Per esempio dalla prima abbiamo notato diverse cose. Tu, Paolo, cos'hai notato da questa prima storia?
M1. Che anche gli animali si possono aiutare insieme.
Ins. **Sì, che anche gli animali possono aiutarsi a vicenda.** Ma questa storia l'hai trovata molto diversa dalle nostre oppure hai notato qualche somiglianza?
M1. No, ci sono alcune somiglianze.
Ins. **Direi che ci sono diverse somiglianze. Sono molto simili ad alcune storie che ci raccontiamo qui in Italia.** E ve la ricordate? Tu Gennaro.
M2. C'è una formica che sta affogando e la colomba la prende e la tira su -
Ins. Sì, e poi che succede Marco?
M3. Che poi la formica lo punge per salvarla.
Ins. **Che nel momento in cui un cacciatore sta per sparare alla colomba, la formica lo punge in modo da fargli sbagliare il tiro.** Abbiamo quindi un caso di -
Bambini: Amicizia.

Una seconda struttura dell'interazione include la formulazione come commento implicito che esprime la posizione dell'agente attraverso un'inferenza, collegata ai turni precedenti ma autonoma nel produrre significati nuovi.
La sequenza 2 si svolge in contesto scolastico simile a quello della sequenza 1, durante un'attività sulla gestione dei conflitti, condotta da un esperto esterno alla scuola. La formulazione dell'esperto (turno 15) è collocata al termine di una discussione nel gruppo di bambini su una decisione da prendere. La formulazione riassume i turni precedenti dei

bambini, suggerendo un'inferenza sulla loro implicazione (evidente nella sospensione finale: *dicono* -). Il suggerimento proietta un dubbio che è sviluppato prontamente da una bambina che coglie nella formulazione la risonanza delle proprie emozioni (turno 16).

S2
Esp: voi avete deciso di fare un cartellone con le immagini (.) ma se io adesso vi chiedessi di inventare una storia, sulla base di queste cosa qua, vi andrebbe di fare?
F1: sì, sì
Esp: a me interessa vedere come arrivate alla decisione -
F1: chi vota per la storia?
Esp: a me interessa sapere come arrivate alla decisione
M1: (.) noi facciamo solo le votazioni
F2: facciamo una leggenda!
M1: facciamo un accordo (.) chi vota per la storia? ((solo Katia alza la mano))
M2 ((deridendo F1)): ah, ah
M1: chi vota "no"? ((quattro mani alzate))
M2: ah, ah, 4 contro 1, [allora: -
F1: [però io voglio fare la storia
F2: e va bene, la facciamo (.) la facciamo -
M1: oh, 'scolta -
Esp ((a F2)): **vince la maggioranza, dicono –**
F3: sì, infatti, dicono così, vince la maggioranza, e prima abbiamo incollato le foto che a qualcuno non andava!

La sequenza 3 riproduce una fase di interazione in un villaggio internazionale che ospita bambini di undici anni (nella trascrizione indicati come maschi o femmine e con la loro nazionalità) e giovani educatori (nella trascrizione indicati nello stesso modo con l'aggiunta della L di "leader", secondo la definizione in uso in questi villaggi) (Baraldi, 2009). La formulazione (turno 9) inferisce l'implicazione del turno precedente: se i bambini si aiutano a vicenda in modo efficace, allora la cooperazione è importante. La formulazione è introdotta da un segnale di conferma (*good*) e dal marcatore di discorso *so*, come nei casi studiati da Hutchby (2007). Si può osservare in senso generale che questo marcatore avvia un'azione autonoma, permettendo anche di costruire coerenza e comprensione nell'interazione (Bolden, 2009). In questo caso, *so* introduce l'inferenza sul significato dei turni precedenti. Il segnale di conferma (*good*) indica la rilevanza della sequenza precedente e il marcatore (*so*) introduce l'inferenza.

S3
LFhgk: so: ple:ase raise your hands if you have something to say about it, ok? We saw that group two won, m:: so: why do you think group two won?
F1bra: because there was (.) was: Paulo he was good in °pone-° put the water in the cup
LFhgk: Bjorn?
M1swe: m: they take only a little water so they wasted less water than us
LFusa: Inga?
F1pol: e: they are good
LFhgk: Brad?
M1usa: they help each other so: they were so: effective
LFhgk: **good (.) so: cooperating is very important**

Nella sequenza 4, al turno 9, la leader italiana traduce per i suoi delegati il turno 8 del leader brasiliano. La traduzione riguarda una parte del turno (in neretto) ed è prodotta come formulazione, introdotta dal segnale di riconoscimento *ok* e dal marcatore *allora* (che equivale a *so*). La formulazione enfatizza la gerarchia tra i leader e i bambini (formulata con "rispetto"), mentre il leader brasiliano nel suo turno sottolinea al contrario il diritto di tutti, adulti e bambini, ad essere ascoltati. La formulazione è qui un'interpretazione che porta a una nuova storia, rovesciata rispetto a quella "originale".

S4
LMbra: (10) ok everybody? Just to remind you that (.) **when (.) an adult or any of you participants raise your arm (we know you have) to say something (.) so (.) other people will respect you (..) and when you also raise an arm and remain in silence and everyone can listen to what you-** […] all right? I'll tell you one more time. A raising hand, a high five means: "ehi people, I want to speak. Please, can you hear me? (..) It is not necessary for us to shout, so we can also be respectful. (..) For the next activities we're gonna have some running games inside the gym (.) because outside it's too hot (03) so we will stay here inside. (..) we don't go outside, because outside (.) it's too hot. The most important for us now is to remain (.) here, inside the gym. Translations.
LFita: **ok**, **allora, prima di tutto ricordatevi che quando un leader alza la mano dobbiamo stare in silenzio, perché significa che vuole dire qualcosa, quindi innanzitutto come segno di rispetto, ok?** Adesso facciamo invece un'attività, un gioco in cui bisognerà correre, stiamo in palestra, non andiamo fuori perché è troppo caldo. Ok?

Una terza struttura dell'interazione include una serie di formulazioni, che operano insieme. La serie di formulazioni può sia verificare la

comprensione dell'agente, sia riassumere i turni degli interlocutori in vista di una conclusione generale.
Nella sequenza 5, la prima formulazione della leader greca (turno 4) è in forma interrogativa ed è una richiesta di chiarimento, attraverso la ripresa del nocciolo del discorso della bambina italiana. La chiusura della prima domanda interna alla formulazione segnala un dubbio interpretativo (*you mean*). La formulazione successiva è suddivisa in due turni (8 e 10), inframmezzati da un segnale di comprensione della bambina italiana. La seconda formulazione (turno 8) completa il chiarimento avviato al turno 4, in cooperazione con le bambine. La terza formulazione (turno 10), introdotta dal marcatore *so* è un'inferenza dell'implicazione dei turni delle bambine. La serie di formulazioni promuove l'affiliazione delle bambine: si produce una sintonizzazione, un "riconoscimento" di affinità, confermato dal fatto che la bambina argentina "riconosce" un senso condiviso nell'interazione (turni 12 e 14)

S5
Fita: yeah but we can do the: (.) they (02) we can ahm no they can: use another, another sense, because the eyes are, must use
LFgre: mm mm
Fita: you must use
LFgre: **because it's night you mean? We must use them?**
Fita: yeah, we must use (both the eyes) (??) you understand?
Farg: no
Fita: but
LFgre: **they cannot see**
Fita: yeah
LFgre: **so they must use another sense**
Fita: *((annuisce))* another sense
Farg: ahh! I understand
Fita: yea. And we can ahm I don't know do: ahm a (?) or (?)
Farg: ah I understand! I understand what she's saying

Nella sequenza 6, la prima formulazione (turno 11) è anticipata da un dubbio della bambina argentina sulla comprensione della leader (turno 8; *do you understand?*), manifestata nonostante i precedenti segnali di incoraggiamento della leader (turni 3, 5). La formulazione, in forma interrogativa, è una richiesta di chiarimento, ma è anche una riparazione del turno 9, nel quale la leader aveva risposto positivamente alla domanda della bambina (*yea, I understood*): è

infatti introdotta dal marcatore *but* che segnala il dubbio sull'affermazione precedente. La formulazione proietta una smentita (turno 13) che conferma la mancata comprensione della leader. La seconda formulazione della leader (turno 14) è introdotta da un segnale di riconoscimento (*yea*) e dal marcatore *so*, è ancora in forma interrogativa e accerta la comprensione. Questa formulazione è confermata e la terza formulazione (turno 16), introdotta ancora dal marcatore *so*, accerta in via definitiva il senso dei turni della bambina argentina, riassumendoli, e rilancia l'interazione, pur rimanendo sospesa a indicare un residuo margine di dubbio. La terza formulazione proietta una nuova e definitiva conferma della bambina argentina. La serie di tre formulazioni, intervallate dai turni della bambina, porta alla condivisione del significato dell'interazione, attraverso un percorso di chiarimento progressivo, punteggiato da tentativi di comprensione.

S6
F1arg: ah for example (…) ah:m he can't see ah we put like an activity that he, he, he (?) can't do, only he can do and I need to (.) to: (.) to cross ahm one place without ahm falling with some objects and the other ahm, the other participants in that group have to tell him what [where he have to go
M1dan: [where he can go
LFgre: mm mmm
F1arg: so that he doesn't fall and in that way they are helping him
LFgre: yea
F1arg: ahm although she can't see
F1ita: ok
F1arg: do you understand?
LFgre: yea, I understood
M1dan: I also [have
LFgre: [**but there, there are gonna be six groups and the six groups are gonna do something at the same time?**
(02)
F1arg: no, no! they are gonna be in different places
LFgre: **yea, so they are gonna be in different places at the same time doing different things?**
F1arg: yes and then they are going to (.) change, like to go around
LFgre: **so every group has to go to every –**
F1arg: every situation yea

Nella sequenza 7, la formulazione in forma interrogativa della leader olandese (turno 46) è introdotta da un segnale di riconoscimento (*ok*) e

dal marcatore *so*. La formulazione è una richiesta di conferma e riceve una conferma indiretta attraverso un commento (turno 47). La seconda formulazione (turno 48), introdotta dal marcatore *but*, mette in discussione la posizione degli interlocutori, discutendo il punto chiarito in base alla prima. La discussione che ne segue segnala una mancata condivisione e porta alla terza formulazione della leader (turno 54), ancora introdotta dal marcatore *but*, che mette nuovamente in discussione la posizione degli interlocutori ed è seguita e rafforzata da uno sviluppo in forma di inferenza, mitigato da una verifica dell'interpretazione (*right?*). La terza formulazione è confermata, sempre in modo indiretto attraverso un commento (turno 55). La serie di formulazioni produce dubbi sul punto di vista degli interlocutori, benché sia mitigata da forme interrogative e da verifiche dell'interpretazione, che non incontrano risposte chiare e dirette, ma commenti.

S7
LFdut: **ok. So you think it's not working?** *((rivolgendosi a Feng))*
Feng: I've just put all together
LFdut: **but you are not (sure) that it will work** *((rivolgendosi a Mdut))*
Mcos: she's just saying that it's not stable
Mdut: it's stable, you can put it
Mcos: no, no
Mdut: if we do more tape we can also can do this
Mcos: we can do –
LFdut: (??) **but maybe you can work on** (.) **on make it** (..) **I think you don't like it because you don't like the tape, right? You don't like [the tape**
Mdut: [it's just ugly

La sequenza 8 si produce in un campo internazionale che coinvolge adolescenti e operatori (Baraldi and Farini, 2009). La prima formulazione dell'operatrice (turno 16), introdotta dal marcatore *so*, riguarda una precedente sequenza piuttosto lunga (qui non riportata) e riassume i punti di vista di due adolescenti che vi hanno contribuito. La seconda formulazione dell'operatrice (turno 25), introdotta da un segnale di riconoscimento (*ok*), riassume la posizione appena espressa da un terzo adolescente (turno 24). L'interruzione di M2ita (turno 26) impedisce all'operatrice di chiudere la formulazione che è quindi completata al turno 27. Al turno 29, dopo un altro tentativo di interruzione, l'operatrice chiede conferma della correttezza della sua formulazione. La seconda formulazione è quindi prodotta in tre turni

(25-27-29) a causa delle interruzioni di Mita2. La serie di formulazioni promuove la partecipazione degli interlocutori attraverso riassunti, con esiti imprevedibili che sono accettati come tali.

S8
OpF1: sorry, sorry, sorry, the other don't (..) **so, M1ita is saying it depends -, if you live in Spain, you are accepted, if you live in England no (..) why Spain and England, sorry? and then M2ita was saying it depends from the state, for example in Spain they are -, it's possible for them to marry, for the people that are homosexual to marry** (03)
M3ita: for me -, the possibility in Spain to get married it isn't means be accepted by the people, I think in English and Spain look homosexual in the same way other people do another way
OpF1: **ok, M3ita then is saying it doesn't really depend on the laws, if I understood well eh, block me if do not, ehm, if -, it doesn't really depend on laws because it can be that it depends also from the people, that live in a country, probably in Spain and in England you can have both behaviour** -
M2ita: sì non c'è micca -
OpF1: **so people that don't accept and people that accept** -
M2ita: sì ma non c'è micca
OpF1: is it right?

Nella sequenza 9, due formulazioni dell'operatrice (turni 3 e 7), introdotte dal marcatore *so,* inferiscono le implicazioni dei turni precedenti (rispettivamente turno 2 e turno 6) e sono entrambe seguite da una seconda parte nello stesso turno: nel primo caso (turno 3) una domanda che invita l'interlocutore a generalizzare le sue considerazioni; nel secondo caso (turno 7), una proposta di generalizzazione da parte dell'operatrice. L'azione dell'operatrice (come si vede anche al turno 1) invita gli interlocutori a generalizzare partendo dalle loro storie e le formulazioni hanno senso in base a questa struttura dell'interazione.

S9
1 OpF1: and how did you fulfil the gaps (...) that you don't have? For example (...) if I imagine to be an illegal immigrant from Mali (...) I mean for example (...) M1alb (...) How did you fulfil the gaps you don't have about a Chinese boy (...) a Chinese son of a Chinese father that hold a successful restaurant? How were you able to fulfil the gaps (...) so to imagine? What did you use in your mind to fulfil? What you didn't know about it?
2 M1alb: I tried to imagine like that because I did myself in that position because one of my cousin is in Sweden and he is Albanian and he owns a middle successful restaurant in Sweden and he told me a little bit what they can do abroad

(..) especially in Sweden and I thought his position, and I compared him (...) his position with the Chinese boy (...) that was the key.
3 OpF1: **so you had something direct (...)** but how you compared the situation, Chinese and Albanian in Sweden for example? What has to do? It has to do something?
4 M1alb: example Albanians are black and in Sweden as he said me they saw you that you don't have (...) not the majority (...) they saw that you don't have for example -
5 M2alb: white skin.
6 M1alb: like this (?) there was some directions I used to fulfil the gaps.
7 OpF1: **so it was the same** (...) that you are different in a foreign country (...) you are a people that is different in a country and you realize that you are different from the other and (...) you can do something but not all the things that the other do -
8 M1alb: not everything (...) not all the things that the others do in their country (...) you have some things.

Una quarta struttura dell'interazione include la formulazione in una sequenza in cui è combinata ad altre tipologie di azione dello stesso agente ed ha significato in questa combinazione.
Nella sequenza 10, la formulazione dell'operatrice (turno 2), che riguarda il turno precedente, è introdotta da un segnale di riconoscimento enfatizzato (*ah ok*) e da un marcatore (*so*) che annuncia un'interpretazione, la cui correttezza è confermata dall'interlocutore (turno 3). La conferma permette all'operatrice di proseguire con una domanda promozionale rivolta al gruppo (turno 4: *were you all agree?*), ancora confermata, e quindi con un segnale di riconoscimento del contributo innovativo di M2fra (turno 8: *ah ah*) che ne proietta la continuazione (turno 9). La formulazione è combinata a segnali di riconoscimento e rilanci della partecipazione: il successo comunicativo della formulazione proietta la continuazione dell'interazione che è ulteriormente promossa da altre azioni dell'operatrice.

S10
Mger: religion shouldn't take decision, political decision
OpF2: **ah ok, so you didn't agree with the (04) with the putting in of the religious theme into this text for all European citizens**
Mger: (yeah)
OpF2: were you all agree?
Adolescenti: yes
Ffra: yeah, but we had to talk a bit because (03) ((chiede qualcosa a M1fra))
M2fra: and we take the topic "united in diversity"

OpF2: ah ah
M2fra: because it's the (??) of the European union and also the motto for the European union, united in diversity, so it's really many nations but we are united, we're together (03) go on somebody (..)

Nella sequenza 11, la formulazione (turno 2) è simile a quella della sequenza 10, ma è introdotta in modo più semplice (*yes, so*) e esplicitamente volta alla verifica dell'interpretazione (*if I got well*). L'uso della formulazione come verifica dell'interpretazione proietta in questo caso un dissenso aperto (turni 5, 7). L'operatrice promuove l'interazione con inviti a continuare, dapprima in modo indiretto (turno 4: *and*), poi in modo mirato (turno 8), infine in modo generalizzato (turno 10).

S11
Mger: when homosexual people educate a small child, this child would be with this, with this ill people, with this are -
OpF1: **yes, so, if I got well, it's (..) ok, it's that if they are educated, if a child is educated in a family of homosexual people they get the same illness?**
Mger: yes
OpF1: and -
M3ita: I think the opposite
OpF1: eh?
M3ita: I think the opposite
OpF1: please, say your opinion
M3ita: if a child is grow up by maybe two men, two people that are looked different by the society, this two people may (..) insegnare è (.) teach to this child to accept all the different people from you and have no prejudice, maybe it's better grow up with two homosexual parents than two normal parents
OpF1: did you understand what M3ita said, everybody understood? (??) you understood as well? ok, and what do you think about it (..) and then I'm very curios as well (..) yes ((vede Fger che vuole parlare))

Nella sequenza 12, la formulazione dell'operatrice (turno 17), introdotta dal marcatore *so*, riassume una precedente conversazione prolungata con alcuni adolescenti (qui non riportata). La formulazione è confermata (turno 18) e l'operatrice rilancia la partecipazione (turno 19) e poi fa da eco (turno 21) a un contributo innovativo (turno 20). In questo caso, la struttura dell'interazione include una sequenza di formulazione, rilancio della partecipazione ed eco.

S12
17 OpF1: **so you could never answer yes (...) or quite never (...) to the question -**
18 Fita: no, never.
19 OpF1: What about you Katrine?
20 Fger: I also like her, because I couldn't move (...) because I can't say yes to the question.
21 OpF1: you can't say yes to the question -

La sequenza 13 riproduce una fase di un'interazione bilingue tra una mediatrice arabofona, un medico e una paziente (Baraldi and Gavioli, 2007). La formulazione della mediatrice si produce in tre turni (21, 23, 25) ed è preceduta da eco che incoraggiano la manifestazione di dubbi e difficoltà da parte della paziente (turni 12 e 16) e che hanno successo, promovendo sintonizzazione e riconoscimento (turni 13, 15, 17, 19). L'incoraggiamento consente alla paziente di narrare le sue emozioni, il cui nocciolo viene successivamente formulato dalla mediatrice attraverso la traduzione. La formulazione inizia con un riassunto del contenuto dei turni della paziente (turno 21), continua con un passaggio verso un'inferenza (turno 23) e si conclude con un passaggio definitivo ad un'inferenza (turno 25), accompagnato da un sorriso che manifesta comprensione. La formulazione in tre turni è sostenuta da segnali di incoraggiamento da parte del medico (turni 22, 24) e proietta un'azione che manifesta il coinvolgimento del medico (turno 26).

S13
DF: Se invece lei vuole che la guardo (.) ok (.) volentieri (.) ho tempo la posso anche controllare
MF: Certo. بتقلك انتي هلئ بشكل عام فيه عندك حاجة مشكلة مضايقتك؟ حاجة والة كلشي يعني بخير ما عندك ايه حاجة؟ لان بتقول اذا عندك ايه مشكلة(.) لانو بالعادة لما تركب الوحدة اللولب(.)تقريبا كل سنة بتعمل الفحوصات
ti dice adesso tu generalmente hai qualcosa, qualche problema che ti disturba? Oppure tutto a posto e non hai niente? Perché dice se non hai nessun problema (.) perché normalmente quando una mette la spirale (.) quasi ogni anno fa i controlli
DF: Perché se no è verso luglio agosto (.) insomma quest'estate
MF: Eh
DF: Dopo un anno
MF: "يعني انتي لما بتركبي بشهر سبعة فيكي تستني لشهر سبعة اخر العام المقبل حتى تعملي "الكونترول السنوي
Cioè se lo metti il mese di Luglio potresti aspettare fino a Luglio dell'anno prossimo per fare il controllo annuo

PF:-هلئ

adesso –

MF: هلئ-

adesso -

13. PF: ما عندي شي
non ho niente
14. MF:
هذا اذا كان ملشي بخير وما عندك مثلا بيهبط منك الدم بزاف وما حاسة حالك ساعتها مثلا بتقول انو ما في حاجة اني اعملك الفحص
questo (.)se tutto va bene e non ti scende tanto la mestruazione e non senti niente (.)allora ti dice che non è necessario il controllo
15. PF:
قضى الشهر وباقي مجتنيش
adesso è un mese e ancora non mi è venuta
16. MF:
شهر ومجتكيش
un mese e non ti è venuta
17. PF:
شهر نيشة
un mese preciso
18. MF:
مش يعني –
non è che –
19. PF:
ما حسيت شي
non ho sentito niente
20. MF:
اه (.) طيب
ah (.) va bene
21. MF: **Dice che dolore (.) qualcosa di strano non c'è (.) dice che sta bene**
22. DF: Ah
23. MF: **Ha le mestruazioni abbondanti (..) l'unica solo cosa forse è per questo che è venuta**
24. DF: Mmh
25. MF: **Che la mestruazione questo mese non è venuta (..) [sorridendo] e lei è un pochino preoccupata**
26. DF: Allora (.) al limite facciamo una cosa (.) le facciamo fare un test di gravidanza

Le formulazioni sono incluse in strutture dell'interazione diverse e di conseguenza hanno significati diversi: possono essere correzioni mitigate (S1), commenti (S2, S3, S4), verifiche della comprensione (S5, S6, S7), riassunti che "muovono" la posizione degli interlocutori (S8, S9), promozioni della partecipazione degli interlocutori (S10, S11, S12, S13). Per comprendere le diverse formulazioni è necessario

analizzare sia l'orientamento ai turni precedenti, sia la proiezione dei turni successivi, nell'interazione complessiva.
Anzitutto, le formulazioni sono precedute e proiettate da sequenze strutturate: intere sequenze diadiche, oppure domande specifiche. Il collegamento con i turni precedenti è frequentemente indicato da segnali di conferma della comprensione e di riconoscimento dell'azione (*ok*, *ah*, *yea* ecc.) e dal marcatore di discorso *so*. L'uso di *so* non è determinante, ma è piuttosto determinato dalla struttura dell'interazione: segue la varietà delle strutture dell'interazione e delle formulazioni, per cui introduce riassunti, come in Hutchby (2007), ma anche commenti e inferenze.
In secondo luogo, le formulazioni sono seguite da seconde parti di turno che ne integrano la funzione, da altre formulazioni che ne precisano il senso, o da altri tipi di turni che ne integrano l'operare nell'interazione.
Infine, i turni degli interlocutori sono fondamentali per generare il senso complessivo dell'interazione che ingloba la formulazione e quindi anche la formulazione stessa. In particolare, le altre azioni che seguono le formulazioni, sono proiettate dai turni degli interlocutori che possono essere sia di allineamento (conformità), sia di "resistenza", spesso come smentita (devianza), sia infine di rilancio dell'interazione (innovazione).

4 La formulazione nel contesto

La varietà delle formulazioni in sequenze strutturate di interazione è legata alla differenziazione dei contesti istituzionali. Il significato della formulazione dipende dai presupposti culturali (Gumperz, 1992) che danno forma a tali contesti e che si manifestano attraverso le aspettative che guidano l'azione dei partecipanti e quindi l'interazione stessa (Baker, 2006; Luhmann, 1984). I presupposti culturali forniscono orientamenti che integrano la singola interazione in sistemi sociali più ampi: nei casi specifici in esame, il sistema educativo e il sistema sanitario.
I presupposti culturali sono indicati nell'interazione attraverso l'uso del linguaggio (Gumperz, 1992): l'analisi delle sequenze presentata nella sezione precedente richiede quindi di essere integrata con l'osservazione dei presupposti culturali delle interazioni. Nelle sequenze si possono infatti osservare gli indicatori delle forme di

aspettative che sono alla base delle interazioni e quindi anche delle formulazioni. In particolare, si possono distinguere tre forme di aspettative che orientano interazioni e formulazioni.
In primo luogo, è possibile osservare interazioni e formulazioni basate su aspettative cognitive, che riguardano la produzione di conoscenza (S1, S9). In questi casi, le formulazioni (e le interazioni) sono orientate alla costruzione di conoscenze a beneficio della formazione dei giovani, oppure di pazienti e medici, dai quali ci si aspetta allineamento.
In secondo luogo, è possibile osservare strutture dell'interazione e formulazioni basate su aspettative normative, che riguardano il rispetto delle regole (S3, S4, S7). In questi casi, le formulazioni (e le interazioni) sono orientate alla fissazione di norme di riferimento per giovani e pazienti, dai quali ci si aspetta allineamento.
In terzo luogo, è possibile osservare strutture dell'interazione e formulazioni basate su aspettative affettive che riguardano l'espressione personale (S2, S5, S6, S8, S10, S11, S12, S13). Le formulazioni (e le interazioni) sono orientate alla promozione della partecipazione attiva di giovani, pazienti e medici: richieste di chiarimenti, manifestazioni di interesse per la comprensione, altre azioni integrate alle formulazioni (eco, domande promozionali, segnali di riconoscimento o incoraggiamento), favoriscono la partecipazione attiva e innovativa degli interlocutori.
Le forme di aspettative permettono di capire i significati sia dell'uso dei segnali e marcatori linguistici, sia dei tipi di formulazione (correzioni mitigate, commenti, verifiche della comprensione, riassunti, promozioni di partecipazione), dei loro collegamenti a turni precedenti e alle seconde parti dei turni e delle loro proiezioni delle sequenze successive.
I diversi tipi di formulazione e le diverse forme di aspettative che le guidano sono correlabili alle opportunità di cambiamento socioculturale offerte dalle interazioni. Le formulazioni possono favorire il cambiamento delle strutture dell'interazione attraverso interpretazioni che aprono possibilità di conferma, devianza o innovazione, permettendo di selezionare positivamente narrazioni e posizioni sociali dei partecipanti. Variazioni nelle narrazioni e nelle posizioni sociali determinano variazioni nell'interazione (Winslade and Monk, 2008). La chiusura a nuove narrazioni e a variazioni nelle posizioni sociali mantiene invece lo status quo. Le formulazioni basate

su aspettative affettive aprono al cambiamento, mentre le formulazioni basate su aspettative cognitive e normative contribuiscono a mantenere lo status quo. In sistemi come quello educativo e quello sanitario, le formulazioni basate su aspettative affettive possono contribuire al cambiamento delle posizioni di giovani e pazienti nell'interazione. Le formulazioni non hanno il potere di avviare analoghi cambiamenti nel sistema sociale complessivo, ma offrono opportunità di variazione locale che possono essere selezionate positivamente in tale sistema.

5 Conclusioni

Lo studio delle formulazioni è di particolare interesse per l'analisi dell'interazione. Per questo studio, è importante integrare due livelli di analisi: 1) l'inclusione della formulazione in una struttura dell'interazione; 2) il collegamento della formulazione ai presupposti culturali del sistema entro il quale l'interazione si produce, ossia alle forme di aspettative che lo caratterizzano.
Questa integrazione porta a osservare le formulazioni non come azioni a se stanti o agganciate ai turni immediatamente circostanti, ma come azioni incluse in un'interazione complessiva guidata da particolari presupposti culturali. Inoltre, permette di spiegare i "comportamenti" variegati di lessico e strutture grammaticali come dipendenti dalle strutture dell'interazione e dai presupposti culturali. Per fare riferimento a Hutchy (2007), né una sequenza QAF o QUAFQ, né un uso del marcatore linguistico *so* possono essere considerati in modo avulso dalla struttura dell'interazione e dai presupposti culturali del sistema sociale che la include. Questi fenomeni linguistici non sono peraltro legati a una determinata struttura dell'interazione o a un determinato sistema sociale, bensì possono essere prodotti in interazioni diversamente strutturate e in base a presupposti culturali diversi. D'altro canto, la formulazione è un fenomeno interessante anche perché può innescare localmente cambiamenti nei presupposti culturali di un sistema sociale. Il cambiamento locale non determina di per sé un cambiamento socioculturale generale, ma crea variazioni che possono essere selezionate in vista di tale cambiamento. Benché sia prematuro trarre conclusioni generali, le formulazioni sembrano essere particolarmente efficaci nell'innescare processi di questo tipo.

Riferimenti bibliografici

Baker, M. (2006) 'Contextualization in translator- and interpreter-mediated events'. *Journal of Pragmatics* 38, pp. 321-337.

Baraldi, C. (2007) La comunicazione istituzionale come dialogo: uno studio in ambito scolastico, in C. Bosisio, B. Cambiaghi, E. Piemontese and F. Santulli (eds), *Atti del 7° Congresso dell'Associazione Italiana di Linguistica Applicata*, Perugia: Guerra, pp. 411-430.

— (ed.) (2009) *Dialogue in intercultural communities. From and educational point of view*. Amsterdam: John Benjamins.

Baraldi, C and F. Farini (2009) *Campi a Monte Sole*. Roma: Carocci.

Baraldi, C. and L. Gavioli (2007) Dialogue interpreting as intercultural mediation. An analysis in healthcare multicultural settings, in M. Grein and E. Weigand (eds), *Dialogue and culture*, Amsterdam: John Benjamins, pp. 155-175.

Bolden, G. (2009) 'Implementing incipient actions: The discourse marker 'so' in English conversation'. *Journal of Pragmatics* 41, pp. 074-998.

Goodwin, C. and J. Heritage (1990) 'Conversation analysis'. *Annual Review of Anthropology* 19, pp. 283-307.

Gumperz, J. (1992) Contextualization and understanding, in A. Duranti and C. Goodwin (eds), *Rethinking context: Language as an interactive phenomenon*, Cambridge: Cambridge University Press, pp. 229-253.

Heritage, J. (1985) Analysing News Interviews: Aspects of the Production of Talk for an Overhearing Audience, in T. Van Dijk (ed.), *Handbook of discourse analysis, Vol. 3. Discourse and dialogue*, London: Academic Press, pp. 95-117.

Hutchby, I. (2007) *The discourse of child counselling*. Amsterdam: John Benjamins.

Luhmann, N. (1984) *Soziale systeme*. Frankfurt a.M.: Suhrkamp.

Mearns, D. and B. Thorne (1999) *Person-centred counselling in action*. London: Sage.

Winslade, J. and G. Monk (2008) *Practicing narrative mediation: Loosening the grip of conflict*. San Francisco: Jossey-Bass.

Il parlato nella storia dell'antropologia statunitense. Un progetto politico

Alessandro Duranti
UCLA University of California at Los Angeles
aduranti@anthro.ucla.edu

1 Introduzione

La discussione di alcune fasi travagliate dello studio delle lingue amerindiane in quanto lingue "parlate" ci offrirà l'occasione di verificare la tesi che l'attenzione al parlato di per sé non garantisce o favorisce un tipo di analisi piuttosto che un'altra, vale a dire, l'attenzione al parlato non ci rende immediatamente teorici del sociale o della cultura. Ciò che conta è piuttosto l'atteggiamento che il ricercatore ha verso la comunicazione parlata e la misura in cui tale atteggiamento lascia spazio al contesto del parlato. Sebbene il parlato sembri offrire l'opportunità di collegare il lavoro del ricercatore (o quantomeno orientarlo) verso il mondo della vita dell'atto di parola e quindi verso i suoi produttori e consumatori, ciò non avviene *naturalmente*. C'è piuttosto bisogno di assumere un particolare tipo di atteggiamento verso l'oggetto di studio, un atteggiamento che definisco "politico" nel senso di un atteggiamento che da una parte è teso alla trasformazione dei pregiudizi linguistici e dall'altra produce una documentazione affidabile e contestabile delle condizioni socio-storiche della vita delle lingue.

2 Il parlato nell'antropologia nordamericana

Nel contesto della professionalizzazione dell'antropologia negli Stati Uniti, il parlato emerge a partire dalla fine del diciannovesimo secolo come la fonte prima di informazioni etnologiche oltre che linguistiche. Come vedremo, esso è visto come un oggetto di studio che è al tempo stesso *prezioso* ed *effimero*. La combinazione di queste due valutazioni contribuisce ad un enorme lavoro di documentazione e di fermento intellettuale che avrà ripercussioni nella trasformazione dell'antropologia statunitense in disciplina accademica di primo piano

nelle scienze sociali e contribuirà alla celebrazione delle culture e delle lingue del mondo al di fuori di schemi evoluzionistici e razzisti. Lo studio delle lingue amerindiane ha avuto un ruolo di primo piano in questa impresa intellettuale e fin dall'inizio di rivendicazione sociale, e quindi politica.

Al tempo stesso nel suo divenire l'antropologia nordamericana, nel provarsi come vera scienza, e nel tentativo di salvare e preservare il prezioso e l'effimero della comunicazione parlata, finirà col fare anche dei gravi danni ad alcuni individui e ad alcuni gruppi nel momento stesso in cui voleva salvarne la tradizione o il *Geist* di herderiana o humboldtiana memoria (Bunzl, 1996), entrando in un circolo vizioso che macchierà la storia della disciplina antropologica fino ai giorni nostri e che porterà negli anni Ottanta alla critica post-coloniale e post-moderna dell'idea stessa di etnografia e partecipazione osservante.

Sebbene questi problemi siano specifici di una disciplina che nasce e si sviluppa in un contesto storico e sociale assai diverso dal come si è sviluppato lo studio del linguaggio e delle lingue storico-naturali in Europa, credo che un'analisi delle condizioni in cui il parlato è stato studiato e valorizzato nell'antropologia statunitense ci possa fornire degli spunti importanti per la considerazioni di situazioni assai diverse, come quelle ad esempio in cui ci si trova a studiare la propria lingua parlata (anziché quella degli altri) e si pensa di poter ignorare le condizione socio-storiche che permettono o meno la documentazione della comunicazione parlata. In altre parole, l'esercizio storico-critico che propongo in questo articolo vuole essere non tanto un esercizio di condanna morale dei nostri (o almeno miei) antenati intellettuali quanto piuttosto un'occasione per capire meglio il tipo di rapporti che in quanto scienziati del parlare possiamo o dobbiamo instaurare sia con il nostro oggetto di studio che con i soggetti che ce lo offrono, consciamente o meno, tramite le loro azioni. O meglio ancora, il mio è un invito a comprendere il modo in cui la costituzione stessa del nostro oggetto di studio partecipa alla costituzione non solo dei nostri soggetti di ricerca ("i parlanti") ma anche alla costituzione di noi stessi come promotori-esecutori di tale ricerca. Tutto questo è estraneo alla linguistica strutturalista o generativa, che è fondamentalmente rimasta acritica verso la propria storia e i propri metodi.

Considero quindi il mio progetto potenzialmente rilevante per delle questioni più generali riguardanti non solo l'analisi linguistica e la

teorizzazione della comunicazione, ma anche il rapporto con gli individui e le comunità di cui studiamo le pratiche comunicative. È quest'ultimo un tema ricorrente dell'antropologia statunitense almeno dalla metà dell'Ottocento – agli inizi del lavoro del Bureau of Ethnology – fino ai giorni nostri. È proprio il contatto diretto con i nativi e la partecipazione alla loro vita sociale che permette all'antropologo un tipo di sapere diverso, per certi aspetti privilegiato, rispetto al sapere di altri scienziati del sociale che si servono di altri metodi di ricerca che li tengono a distanza fisica e affettiva dai loro soggetti. Ed è anche il contatto diretto, l'attingere conoscenze da esseri viventi che rende il lavoro dell'antropologo delicato e pericoloso, con possibili conseguenze psicologiche, fisiche, emotive, e politiche per chi ne è partecipe. È all'interno di questa tradizione piena di rischi, creatività, sacrifici, ed errori che si situa l'antropologia linguistica contemporanea. È mia convinzione che la conoscenza di questo passato non può che giovare al nostro futuro.

3 Boas e il ruolo delle lingue amerindiane nella nascita dell'antropologia statunitense

L'antropologia statunitense ci offre l'opportunità di pensare al linguaggio parlato in una prospettiva storica e culturale che ci stimola a prendere in considerazione il ruolo delle innovazioni tecnologiche (e.g. Brady, 1999) e dell'ideologia (Bauman and Briggs, 2003) nella concettualizzazione dell'oggetto di ricerca. Diversamente da quella europea, l'antropologia che si pratica negli Stati Uniti nasce alla fine del diciannovesimo secolo non solo come progetto di studio delle culture ma anche come progetto di studio di lingue native d'America, lingue che fino ad allora non erano state documentate o erano state documentate male (Boas, 1889). Il giovane Franz Boas, che in pochi anni si trasforma da geografo in etnologo e linguista, viene sovvenzionato dal direttore del Bureau of Ethnology[1], John W. Powell, col preciso scopo di descrivere le lingue degli indiani d'America e i loro rapporti genetici (McGee, 1897)[2]. Questo lavoro

[1] Il Bureau of Ethnology fu fondato nel 1879 con fondi del governo federale USA. Nel 1892 fu rinominato "Bureau of American Ethnology" (Conn, 2006: 176).

[2] Statesmen and administrative officers concerned with placing the Indians on reservations felt the need of a practical classification of the Indian tribes under

sarebbe dovuto servire al governo federale americano per classificare e gestire politicamente e militarmente le tribù degli indiani d'America (ad esempio per decidere quali gruppi mettere nella stessa riserva). Come sappiamo, Boas si guardò bene dal proporre schemi genetici e sviluppò invece una visione "diffusionista" dei rapporti tra lingue e tra culture (bisognerà aspettare i lavori del suo geniale studente Edward Sapir per arrivare a una rivalutazione della linguistica storica e comparativa e ad una riformulazione della classificazione genetica proposta da Powell).

Anche se l'importanza delle lingue delle popolazioni native dell'America del Nord già si trova nella pratica del Bureau of Ethnology diretto da Powell (McGee, 1897)[3], Boas ebbe un ruolo fondamentale nel promuovere e teorizzare l'antropologia come disciplina che doveva comprendere non solo lo studio delle caratteristiche fisiche dei vari gruppi sociali, e i loro costumi, ma anche le loro lingue. I dati linguistici per Boas era ancora più preziosi di altre manifestazioni culturali grazie alla loro natura non-conscia e quindi immune da quelle che lui chiamava generalizzazioni "secondarie" ovvero razionalizzazioni dei nativi senza prove empiriche.

La documentazione e comprensione della complessa morfo-sintassi delle lingue amerindiane era per Boas un modo di demolire l'evoluzionismo culturale predominante nel diciannovesimo secolo. Ispirato dalle idee liberali dei moti sociali europei del 1848, con un'esperienza alle spalle di lotta, anche fisica (Liss, 1996: 168-9), contro l'antisemitismo dei suoi compagni d'università, come linguista Boas fin dall'inizio lavora per un cosmopolitismo in cui le lingue e le tradizioni degli indiani d'America possano essere riconosciute come

---

which they might be arranged in amicable groups; this need was urged on Major Powell while Director of the Rocky Mountain Survey, and the anthropological researches of the survey were bent to meet it; and when provision was made for continuing the work it was understood that the primary duty of the new bureau should be the classification of the Indian tribes for practical as well as for scientific purposes (McGee, 1897: 373)

[3] Foremost among these [matters], not only as the basis of all the others but in immediate important, is language, including speech and the germ of writing; second in order of development and importance come the arts, esthetic and industrial; next in order are institutions; and perhaps youngest in origin and most interesting to thoughtful investigators are beliefs (McGee, 1897: 375)

di pari valore alle lingue e le tradizioni europee (Lettera di Boas a Bolton Landing, July 24, 1905, in Stocking, 1974: 122-3; Boas, 1911). È questo un programma in cui l'antropologia, o meglio quella parte dell'antropologia che si chiamava allora "ethnology" può essere di aiuto ad una visione più "oggettiva" di noi stessi (Boas in Stocking, 1974: 71) e quindi, diremo noi, meno etno-centrica. Ad esempio, il 15 febbraio 1905, nello scrivere a W. Thalbitzer per invitarlo a contribuire un suo capitolo sull'eschimese, Boas consiglia di rimanere "puramente analitico, vale a dire, di evitare il più possibile il punto di vista delle lingue indo-europee" («purely analytical; that is to say, to keep out the point of view of Indo-European languages as thoroughly as possible») (Stocking, 1974: 178). In questo Boas continuava la tradizione già stabilita all'interno del Bureau of Ethnology, come attesta la storia del Bureau proposta da McGee (1897)[4].

4 Il paradosso dell'orale che diventa scritto

Ma è proprio l'atteggiamento progressista e romantico della sua generazione di intellettuali eredi di Herder, che tende una prima trappola a Boas. Elevare lo status delle culture amerindiane sia nel mondo accademico che nell'opinione pubblica vuol dire per Boas fissare le lingue in modo da costituirne una tradizione "letteraria" comparabile a quella delle lingue classiche.

Poiché le lingue degli amerindiani non erano scritte, fin dall'inizio il progetto di documentazione di tali lingue fu un progetto di documentazione di lingue che esistevano *solo* come lingue parlate, ma che, per poter essere studiate e conservate come patrimonio dell'umanità, dovevano essere trasformate, nel più breve tempo possibile, in lingue scritte. Nel far questo, l'etnologo-linguista si impegna alla ricerca delle culture e delle lingue "autentiche" e non si rende conto del fatto che automaticamente ignora le situazioni di contatto, il sincretismo dilagante (Bauman and Briggs, 2003; Hill, 2001; Hill and Hill, 1986) ed entra come protagonista in un processo di acculturazione in cui la lingua in quanto rappresentazione della

[4] McGee (1897: 379-80) cita le istruzioni date già nel 1852 dal Prof. William W. Turner sul bollettino della Smithsonian Institution a chi avrebbe richiesto la collaborazione di linguisti: «Let the writer … describe the particular language under consideration; let all fanciful comparison with Hebrew, Greek, etc., be excluded».

cultura viene contemporaneamente separata da essa. Nel trascriverle, Boas congela le lingue amerindiane nel tempo e nello spazio.
Ma la scrittura non "decontestualizza" il linguaggio parlato – caso mai lo ri-contestualizza. Prendendo in prestito i termini usati da Walter Benjamin (2000) per l'opera d'arte nell'era della sua riproduzione meccanica, potremmo dire che la scrittura, come la fotografia rispetto ai capolavori d'arte, rimuove l'*aura*, l'*hic et nunc* del parlare, ovvero il suo carattere indicale (o indessicale). La scrittura agisce sull'autenticità-unicità del dire tramite una trasformazione della temporalità dell'esperienza di decodificazione. Nel caso di Boas e degli altri ricercatori del Bureau of Ethnology, la trasformazione del parlato in scritto è accompagnata non solo da un'esigenza di documentazione accurata (e non quindi da dilettanti) ma anche da un senso di urgenza dovuto ad una situazione storica e sociale molto drammatica, ovvero la constatata decimazione delle popolazioni indigene ed il rapido processo di acculturazione in corso e quindi la conseguente perdita della competenza linguistica da parte delle nuove generazioni sottoposte a dominio militare, a violenza da parte dei colonizzatori bianchi, e a contatto linguistico tra di loro e con le lingue indoeuropee dei coloni. Questo stesso senso di urgenza era la motivazione di fondo di quella che più tardi fu chiamata "salvage anthropology" ("antropologia di salvataggio"), un'impresa che richiedeva continui sovvenzionamenti da enti federali (come il Bureau of Ethnology, poi ribattezzato Bureau of American Ethnology), o da facoltosi sponsor.
Ad esempio, come documentato presso la Bancroft Library alla University of California, Berkeley, nel 1901 Boas scriveva a Zelia Nuttall, una rinomata viaggiatrice, poliglotta, archeologa e studiosa di culture pre-colombiane in Messico (sposata ad un esploratore francese), chiedendole di inoltrare a sua volta alla facoltosa Sig.ra Phoebe Hearst una richiesta per fondi per lo studio delle lingue amerindiane della California. In questa lettera, Boas parla della necessità di raccogliere materiali sul gran numero di lingue della California "delle quali sappiamo praticamente nulla" e che hanno bisogno di essere documentate da studenti capaci, che sappiano di fonetica e di lingue amerindiane, tra i quali ci sarà il giovane Alfred Kroeber, che aveva frequentato i corsi di Boas alla Columbia University alla fine del secolo e che fonderà il dipartimento di antropologia a Berkeley. Ed infatti fu proprio la Sig.ra Hearst a pagare

per i primi cinque anni lo stipendio di Kroeber all'University of California (Hearst pagò anche per lo stipendio di P.E. Goddard, di fatto finanziando la costituzione del dipartimento d'antropologia) (Steward 1961).
La lettera di Boas merita di essere citata quasi per intero perché la formulazione della sua richiesta ci permetterà alcune osservazioni[5].

New York, April 11, 1901.
Mrs. Zelia Nuttall, The Albany, Washington, D.C.
My dear Mrs. Nuttall,--
[…]
I wish to interest you in one subject to which I have been devoting a great deal of time and energy during the last few years. You know that in California we have an enormous mass of Indian tribes and languages about which we know practically nothing. There are a few missionary grammars and quite a number of short vocabularies, which, however, do not amount to very much. You are aware that all these tribes are on the verge of extinction, and that it is only a question of a very few years when their languages, and with them their traditions and the records of their customs, will have disappeared. For this reason I have been exceedingly anxious to train men and to raise money to carry on work among these tribes. I have been somewhat successful in this respect, and during the past three years I have been able to interest some friends of science in New York in this subject, and I have three trained philologists working on the subject; but if we are to save any considerable amount of the information which is now vanishing day by day, it is necessary that the available funds should be much greater than they are at the present time. I am now in a position to spend approximately $3000 a year in California, but the work cannot be done effectively with that amount of money. I need at least $3000 a year more to cover both the expense of collecting information and of publishing the same. I believe you will agree with me in your view of the urgency of this matter.
I did not venture to push this matter strongly until three years ago, for the reason that the number of trained students is so exceedingly limited. To place work of this kind in the hands of the inexperienced man would be disastrous. A student who is to make collections of this kind must have a thorough knowledge of American phonetics and of the structures of American languages all over the continent in order to reach good results.
[…]

[5] Il testo dell'intera lettera si trova sul sito dal titolo "Foundations of Anthropology at the University of California", al seguente indirizzo: http://bancroft.berkeley.edu/Exhibits/anthro/3founders2_nuttall.html

It occurred to me that you may be in a position to present this important subject to Mrs. Hearst, who is doing so much for the interests of science in California, and that she might perhaps become interested in its execution.
[…]
I hope that I shall lave an opportunity to speak to you in regard to this matter next week; and if you think fit, I shall be indebted to you if you will bring the matter to the attention of Mrs. Hearst, and if I could have an opportunity to explain somewhat more fully the object for which we are working.
Yours very sincerely,
Franz Boas.

5 La fonetica come strumento e gli strumenti meccanici: per far fronte all'urgenza

La prima osservazione sulla lettera di Boas a Zelia Nuttall è che la fonetica è definita da Boas come un *mezzo* e non come un'attività o disciplina fine a se stessa. Questo non sorprende troppo visto che Boas non sembrava avere interessi per la teoria fonologica. Ad esempio, stando alla ricostruzione di Judith Berman (1994: 493), nel 1900 Boas distingueva nella sua trascrizione del Kwakiutl (Kwak'wala) 19 segmenti vocalici, laddove avrebbe avuto senso pensare in termini di varianti di 3 o di 5 segmenti di base. A sua volta la documentazione di queste lingue è sia fine che mezzo. È fine perché la lingua rende possibili miti, leggende, e sistemi di classificazione; ma è anche mezzo per poter avere informazioni su molti aspetti delle culture amerindiane. I testi raccolti da Boas, dai suoi studenti, e dai loro collaboratori nativi (e.g. Gorge Hunt), costituiscono la base "oggettiva" su cui costruire descrizioni etnografiche. L'introduzione del lavoro sul terreno per periodi più lunghi (da uno ai due anni) nell'antropologia sociale britannica (a partire da Malinowski) e poi nell'antropologia culturale nordamericana eliminerà questa pratica dei testi scritti. Gli antropologi in quanto etnografi avranno modo di descrivere loro stessi quello che osservano tramite la pratica delle "fieldnotes" (Sanjek, 1990), ma i lettori avranno resoconti sempre filtrati dal racconto non verificabile degli osservatori.
La trascrizione fonetica come strumento di conoscenza fa parte di una serie di strumenti, alcuni dei quali materiali, meccanici, e con i quali è in rapporto di complementarietà o, in certi casi, d'opposizione se non addirittura di competizione. Uno dei più importanti strumenti è il

fonografo inventato da Edison (e costruito da John Kreusi) e messo in commercio nel 1878. Boas fu uno dei primi ad utilizzarlo per il lavoro di documentazione etnografica. La storia recentemente raccontata da Erika Brady nel libro *A Spiral Way: How the Phonograph Changed Ethnography* parte dalla questione del ruolo del fonografo e della sua tecnologia nella costituzione del contenuto di quello che fu registrato. Va qui ricordato che il "contenuto" sono le *voci* dei nativi.

È interessante che Boas incide soprattutto canto e poco linguaggio parlato. Il catalogo preparato da Dorothy Sarah Lee (1979) delle registrazioni sonore della musica e "dati orali" dei nativi nord americani dal 1893 al 1976 elenca più di dieci ore di registrazioni su cilindri di cera attribuite a Boas di cui solo una minima parte (non si specifica la durata) è descritta come "oral data"; il resto è tutto materiale cantato che copre gli argomenti più diversi, da ninne-nanne a danze sciamaniche, a storie di orsi, anitre, spiriti, salmoni, e guerre (Wickwire, 2001). Questo ci fa capire che il parlato veniva visto da Boas come qualcosa che poteva essere trascritto più facilmente e più rapidamente del cantato. Si poteva chiedere a chi parlava per raccontare una storia di rallentare, fermarsi, ricominciare da un certo punto, ma questo era visto come più difficile da fare per il cantato, anche se non era facile convincere i nativi a esprimersi nel cono del fonografo.

I vantaggi o svantaggi del fonografo erano fonti di continue discussioni e nei pochi documenti rimasti troviamo sia sostenitori che detrattori, tra questi ultima la maggior parte degli (etno-)musicologi britannici che consideravano il fonografo inferiore all'orecchio esperto di chi è stato educato alla trascrizione musicale (si veda la corrispondenza tra Percy Grainger, un entusiasta del fonografo, e Cecil Sharp, un suo critico – Grainger, 1908; Yates, 1982).

Ma altri si entusiasmarono all'uso del fonografo per la registrazione delle lingue amerindiane. Uno degli entusiasti fu J. Walter Fewkes (1850-1930), uno zoologo diventato poi studioso di tradizioni native e successore di Powell come direttore del Bureau of (American) Ethnology (un posto rifiutato da Boas). Fewkes nel 1890 usa un fonografo per registrare la musica degli Zuñi e poi scrive dell'importanza dello strumento per la documentazione e preservazione delle lingue amerindiane (1890a). Come nel caso di Boas, Fewkes parla dell'urgenza dell'impresa:

When one considers the changes which yearly come to the Indians, and the probability that in a few years many of their customs will be greatly modified or disappear forever, the necessity for immediate preservation of their songs and rituals is imperative. (Fewkes 1890a: 1095)

Fewkes è alla ricerca di metodi scientifici che producano dei dati il più accurati possibili.

In order that folk-lore, as far as applicable to aboriginal races, may be placed on a scientific basis, an accurate record of the story as told by the reciter is necessary. This can be accomplished by the use of the phonograph, and the records thus made can be indefinitely preserved. (Fewkes 1890b: 496)

Ma Fewkes, nel suo essere scienziato fino in fondo si preoccupa delle condizioni sperimentali e quindi si interessa anche alla reazione dei nativi: «The great difficulty in getting them to repeat their sacred songs and prayers does not come so much from their fear of the instrument as of secularizing what is sacred to them» (1890c: 691).
Molti anni più tardi, nel riproporre la storia del fonografo per l'antropologia, Erika Brady (1999) racconterà dell'inversione di senso del suo uso ovvero dei pellegrinaggi degli amerindiani alla Smithsonian Institution per ascoltare le voci dei loro padri o nonni.

6 Tecnologie della conoscenza

La definizione stessa di "oggetto d'analisi" presuppone quindi una *tecnologia della conoscenza* che comprende ma non si esaurisce in particolari tecniche di trascrizione, come il sistema fonetico basato sulla scrittura alfabetica, e il fonografo appena citato, al quale se ne possono aggiungere altri, compresa la macchina fotografica, la cinepresa, il registratore a nastro e portatile, il video registratore, il computer, ecc. Come ha dimostrato la storia dell'antropologia e anche quella della linguistica, questi strumenti sono stati usati ma sempre poco rispetto al loro potenziale. La storia della diffidenza verso il fonografo si è ripetuta con la video registrazione negli ultimi trenta anni.
Ciascuno strumento a sua volta presuppone, di solito in modo implicito, delle proprietà che emergono solo nell'uso concreto da parte di utenti particolari per particolari scopi, spesso diversi da quelli per i

quali tali macchine sono state costruite o messe in commercio. Le proprietà stesse degli strumenti adottati dai ricercatori sono sempre dunque proprietà interattive che vengono o meno scoperte e utilizzate a seconda delle condizioni socio-storiche in cui la documentazione viene fatta. Il senso di oggetto prezioso ed effimero che si dà alla voce dei nativi americani ha un ruolo fondamentale nel costituire (in un senso fenomenologico e, aggiungerei, internazionale) la voce come specimen da museo.
I cilindri di cera su cui erano state incise le voci degli indiani d'America venivano mandati e preservati alla Smithsonian Institution, insieme ai crani o ai cervelli sotto spirito degli indiani morti. Ma anche gli Indiani d'America vivi erano ospitati nei musei. Un caso famoso è quello degli Inuit portati nel 1897 a New York dall'esploratore e speculatore Robert Peary, trafficante-importatore di pelli, avorio, e meteoriti trovate in Greenland, autore tra l'altro di un falso arrivo al Polo Nord (prima avvalorato dalla National Geography Society e poi negato, 50 anni più tardi). Solo uno dei sei Inuit sopravvive, il piccolo Minik a cui si nasconde per molti anni che le ossa del padre sono conservate nel museo (gli avevano fatto partecipare ad un finto funerale). Boas stesso, allora collaboratore dell'American Museum of Natural History, aveva chiesto a Peary di portare un Inuit di mezz'età.

I beg to suggest to you that if you are certain of revisiting North Greenland next summer, it would be of the very greatest value if you should be able to bring a middle-aged Eskimo to stay here over winter. This would enable us to obtain leisurely certain information which will be of the greatest scientific importance (citato da Harper, 2002: 39)

Peary che vuole far piacere al Museo a cui vende le sue scoperte ed in particolare ad uno dei patroni del museo, suo sovvenzionatore, invece di portarne uno di Inuit, ne porta sei. Appena arrivati a New York, i sei Inuit sono tenuti a bordo della nave per due giorni mentre 30,000 newyorkesi pagano 25 centesimi per vederli insieme alle due meteoriti che Peary aveva portato (erano stati gli eschimesi che gliele avevano fatto "scoprire") (Harper, 2002).

Figura 1: Quattro dei sei Inuit portati a New York da Peary nel 1897: Nuktaq, Uisaakavsak, Minik, e Qisuk. (Peabody Museum of Archeologo and Ethnology, Harvard University, riprodotta in Harper, 2002: 42)

Kroeber, ventunenne, viene incaricato da Boas di studiare gli Inuit. Lo farà e ne scriverà sul *Bulletin of the American Museum of Natural History* e nella rivista *Journal of American Folk-Lore* (Kroeber, 1899a; 1899b; 1899c). Kroeber scrive anche della reazione di uno dei sei, Nuktaq, alla morte di un altro, Atangana, e, secondo la ricostruzione di Harper (2002: 43), sembra rispettoso della loro privacy, come dimostra il fatto che non va al funerale e si serve della testimonianza di un'altra persona, Esther Enutseak del Labrador che parlava inuit. Eppure nel 1916 Kroeber stesso, pur essendosi originariamente opposto all'autopsia dell'ultimo indiano yahi, da lui chiamato "Ishi", finirà col mandarne il cervello alla Smithsonian Institution. Ci vorrà quasi un secolo prima che un'antropologa nel dipartimento fondato da Kroeber, Nancy Scheper-Hughes (2001), avrà il coraggio di raccontare la storia di Ishi e del tormento post-mortem di Kroeber che, apparentemente depresso, si rifiutava di parlarne.

7 Implicazioni e conseguenze al livello istituzionale

Che cosa impariamo da queste storie tragiche di incontri difficili, drammatici e traumatici tra scienziati dell'umano e le genti le cui voci raccontavano il loro passato? Quanto è cambiato dai tempi di Boas e di Kroeber?

A queste domande si possono dare diverse risposte. Il problema dei rapporti con i nativi è diventato con gli anni un tema centrale dell'antropologia nordamericana. In particolare l'antropologia degli anni Ottanta è stata contrassegnata dall'analisi del lavoro sul terreno (*fieldwork*), del rapporto con i nostri "soggetti", e delle tecniche di documentazione. Ne è un esempio importante il libro *Writing Culture* curato da James Clifford e Gorge Marcus (1986), pubblicato in italiano da Meltemi (*Scrivere le culture*). In antropologia linguistica gli studi sull'ideologia linguistica oltre a discutere le teorizzazioni native su lingua e sul suo uso hanno anche riesaminato i modelli occidentali del rapporto tra le lingue, ad esempio l'uso che s'è fatto delle lingue per l'identificazione di unità etniche o possibili future nazioni (Irvine and Gal, 2000). Va qui inoltre ricordato che fin dagli anni Sessanta la sociolinguistica di William Labov è stata un esempio di linguistica progressista, interessata al riconoscimento civile dei dialetti non-standard dell'inglese (si pensi al famoso articolo The Logic of Non-Standard English; Labov, 1969). La sociolinguistica degli anni Novanta ha aggiunto il punto dei vista dei nativi, visto che, come ha dimostrato Marcyliena Morgan (1994), per molti Afro-Americani l'uso del Black English o African American English Vernacular (AAVE) ha connotazioni sociali storiche sia positive che negative, a seconda del contesto. Visto che l'emancipazione sociale viene associata dalla maggior parte degli afro-americani anche alla padronanza dello standard, la dinamica standard-dialetto è complessa e variabile all'interno della comunità linguistica afro-americana. Negli ultimi anni, il successo globale della musica hip hop ha in certi casi sovvertito l'egemonia dello standard rispetto al dialetto, soprattutto nella propagazione dell'importanza dello stile, delle strofe poetiche, e del virtuosismo verbale che ha tolto il primato dell'attenzione alla pronuncia e quindi alla fonologia (Alim, 2008; Alim and Pennycook, 2008).

Per ragioni esterne all'uso delle culture e delle lingue, vale a dire in seguito a sbagli e soprusi di alcuni esperimenti di ricerca medica, da

diversi anni lo stato federale statunitense ha imposto regole di controllo sulla ricerca che coinvolge le persone. Negli Stati Uniti, le università sono state costrette a stabilire dei protocolli severi sul come si interagisce con soggetti di ricerca umani. Tutti i ricercatori – sia docenti che studenti – per poter indagare le pratiche culturali tramite lo studio delle interazioni umane devono sottoporre le loro proposte al vaglio di commissioni di esperti che valutano i metodi e gli scopi di ciascuna proposta, chiedendo trasparenza e avendo l'autorità di bloccare i finanziamenti laddove si riscontra la possibilità di violazioni di norme etiche o la possibilità di danno fisico o psicologico. Particolare attenzione viene data agli individui e gruppi "a rischio," vale a dire alla ricerca sui minori, le persone anziane, i malati, e i carcerati. L'American Anthropological Association, che conta più di diecimila membri, ha un comitato sull'etica e un altro sui diritti umani.

8 Conclusioni

Sono partito da una domanda sul potere di un fenomeno come il "parlato" di favorire o meno certi tipi di analisi linguistica piuttosto che altri. Per contestualizzare questa domanda, ho presentato una brevissima storia del modo in cui il parlato delle lingue amerindiane divenne un tema centrale della ricerca antropologica e di fatto il punto di contatto dei bianchi con le culture delle popolazioni indigene dell'America del nord. In questa storia ho individuato alcuni fattori ideologici e condizioni storico-politiche particolari che portarono allo sviluppo di alcuni metodi particolari per "fissare" le voci degli indiani d'America in testi scritti o in cilindri di cera incisi dal fonografo. Questo ha portato ad un'ulteriore conferma di temi e dubbi già trattati nella letteratura sui rapporti tra i bianchi e le popolazioni indigene dell'America del Nord. È facile, con il senno del poi, accusare i nostri antenati di essere insensibili o addirittura razzisti, anche quando, nel caso di Boas, Fewkes, e Kroeber (diversamente dall'esploratore Peary) le loro ricerche e metodi erano motivati da buone intenzioni come ad esempio la celebrazione delle culture indigene ed un fondamentale anti-evoluzionismo sociale.

È più difficile capire come la nostra situazione attuale di ricercatori si situi anch'essa all'interno di logiche di potere, di fraintendimenti, di sfruttamento, di guerre, e di rivendicazioni sociali. Ancora più difficile

è riuscire a far funzionare la nostra riflessione critica quando ci troviamo a studiare il parlato di persone a noi vicine fisicamente o socialmente, o quando i nostri soggetti parlano in modi simili ai nostri e diamo per scontata la conoscenza del contesto socio-culturale della loro comunicazione, quando cioè per dirla con Susan Philips (1983), la nostra cultura diventa "invisibile".

Ma la storia del passato può aiutarci a ritrovare la consapevolezza del fatto che i soprusi e i fraintendimenti sono sempre ripetibili nel tempo e nello spazio, anche se in gradi e modi diversi, e che è l'atteggiamento che abbiamo rispetto al nostro oggetto di studio ad avere un ruolo fondamentale nel nostro modo di fare scienza. È il nostro atteggiamento che ci prepara o meno ai modi di essere "altri" o rende tali i modi di essere (e quindi anche di parlare) familiari e quindi pieni di significati socio-culturali e politici nascosti.

Nell'usare il termine "atteggiamento" per il nostro modo di porsi rispetto al fenomeno della comunicazione parlata mi rifaccio alla traduzione italiana (di Enrico Filippini) del termine *Einstellung* usato da Edmund Husserl nel suo *Ideen* (Husserl, 1913) (tale traduzione corrisponde a due traduzioni inglesi, la prima "standpoint" e la seconda, più recenti, "attitude"). Husserl usa *Einstellung* in combinazione con diversi aggettivi, ciascuno dei quali definisce un tipo diverso di rapportarsi al mondo circostante (*Umwelt*) o particolari elementi di tale mondo[6]. Mi interessa in particolare quello che Husserl chiama *die natürliche Einstellung* "l'atteggiamento naturale" – che è tutto meno che "naturale", e cioè è fondamentalmente "culturale", visto che è il nostro modo di essere nella vita di tutti i giorni, piena di valori dati implicitamente a cose, parole, persone, rapporti ed eventi.

Da Husserl impariamo anche che una delle capacità degli umani è quella di poter cambiare il nostro atteggiamento verso aspetti particolari del mondo circostante, compresi gli atti di parola nostri o di altri. Per dirla alla Wittgenstein, noi possiamo cambiare l'aspetto di quello che vediamo o ascoltiamo. Il progetto che ci troviamo a dover affrontare è allora: *come* ascoltare la comunicazione parlata? E quale tipo di "ascoltare come" rende possibile un tipo di documentazione piuttosto che un altro?

[6] Lo stesso termine, *Einstellung*, è usato da Wittgenstein (1958) per spiegare il concetto di "aspetto" (*Aspekt*) o del "vedere come".

Il problema allora per una linguistica che sia una "linguistica della prassi" (Duranti, 1988) è di definire come possiamo incorporare anziché ignorare o contenere le condizioni storiche, sociali e culturali che il parlato ci offre in quanto fenomeno che di sua natura richiama, evoca, e sfrutta il contesto d'uso del linguaggio in generale e della lingua (o lingue) in particolare. Uno studio del parlato che sia allora uno studio storico-pratico deve aprirsi anziché chiudersi alla sfida presentata da un fenomeno che è di sua natura spurio e socialmente complesso; un fenomeno che rende vulnerabili i parlanti, gli ascoltatori e chi li studia.

Piuttosto che "naturalizzare" il parlato per renderlo un oggetto di scienza, come fa spesso la linguistica, oppure invece di limitarsi a renderlo testo separato, perché di fatto separabile, dal suo contesto socio-storico, il nostro compito di linguisti della prassi è quello di far sì che esso ci sveli una ad una le tracce dell'umano che esso raccoglie come prodotto di una storia che è sempre sia individuale che sociale. È questo che io vedo come il compito di quella che ho chiamato *etnopragmatica* (Duranti, 2007), ovvero una pragmatica che sia coinvolta in modo critico, e quindi anche politico in senso lato, con le teorie locali del parlare e della sua forza. È questa l'eredità che ci viene sia dai successi che dagli errori di Boas e dei suoi studenti.

Riferimenti bibliografici

Alim, H. S. (2004) Hip Hop Nation Language, in E. Finegan and J.R. Rickford (ed.) *Language in the USA: Themes for the Twenty-first Century*. New York: Cambridge University Press, pp. 387-409.

Alim, H. S. and A. Pennycook (eds) (2008) *Global Linguistic Flows: Hip Hop Cultures, Youth Identities, and the Politics of Language*. London: Routledge.

Bauman, R., and C.L. Briggs (2003) *Voices of Modernity: Language Ideologies and the Politics of Inequality*. Cambridge: Cambridge University Press.

Benjamin, W. (2000) *L'opera d'arte nell'epoca della sua riproducibilità tecnica*. Torino: Einaudi [Prefazione di Cesare Cases. Traduzione di Enrico Filippini].

Berman, J. (1994) 'George Hunt and the Kwak'wala Texts'. *Anthropological Linguistics*, 36, pp. 483-514.

Boas, F. (1889) 'On Alternating Sounds'. *American Anthropologist*, 2 (o.s.) pp. 47-53.
- (1911) Introduction, in F. Boas (ed.) *Handbook of American Indian Languages*, vol. BAE-B 40, Part I., Washington, D.C.: Smithsonian Institution and Bureau of American Ethnology.

Brady, E. (1999) *A Spiral Way: How the Phonograph Changed Ethnography*. Jackson, MS: University Press of Mississippi.

Bunzl, M. (1996) Franz Boas and the Humboldtian Tradition, in G. W. Stocking, Jr. (ed.) *Volksgeist as Method and Ethic. Essays on Boasian Ethnography and the German Anthropological Tradition*, London: The University of Wisconsin Press, pp. 17-78.

Clifford, J. and G. E. Marcus (1986) *Writing Culture: The Poetics and Politics of Ethnography*. Berkeley: University of California Press [trad. it. di P. Vereni (1997) *Scrivere le culture: Poetiche e politiche in etnografia*, Roma: Meltemi].

Conn, S. (2006) *History's Shadow. Native Americans and Historical Consciousness in the Nineteenth Century*. Chicago: The University of Chicago Press.

Duranti, A. (1988) The Ethnography of Speaking: Toward a Linguistics of the Praxis, in F.J. Newmeyer (ed.) *Linguistics: The Cambridge Survey.* Vol. IV: *Language: The Socio-Cultural Context*., Cambridge: Cambridge University Press, pp. 210-228 [trad. it. di B. Turchetta, Etnografia del parlato: per una linguistica della prassi, in B. Turchetta (ed.) *Introduzione alla linguistica antropologica*, Milano: Mursia, pp. 157-80].
- (2007) *Etnopragmatica. La forza nel parlare*. Roma: Carocci.

Fewkes, J. W. (1890a) 'Additional Studies of Zuni Songs and Rituals with the Phonograph'. *American Naturalist*, 24: pp. 1094-98.
- (1890b) 'On the Use of the Phonograph in the Study of the Languages of American Indians'. *The American Naturalist*, 24, pp. 495-9.
- (1890c) 'On the Use of the Phonograph Among the Zuni Indians'. *The American Naturalist*, 24, pp. 687-91.

Grainger, P. (1908) 'Collecting with the Phonograph'. *Journal of the Folk Song Society*, 12, pp. 147-69.

Harper, K. (2002) 'The Minik Affair: The Role of the American Museum of Natural History'. *Polar Geography*, 26, pp. 39-52.

Hill, J. H. (2001) Sincretismo/syncretism, in A. Duranti (ed.) *Culture*

e discorso: un lessico per le scienze umane, Roma: Meltemi [trad. it. di Antonio Perri].

Hill, J. H. and K.C. Hill. (1986) *Speaking Mexicano: Dynamics of a Syncretic Language in Central Mexico*. Tucson: University of Arizona Press.

Husserl, E. (1913) 'Ideen zu einer reinen Phänomenologie und phänomenologischen Philosophie'. *Jahrbuch für Philosophie und phänomenologische Forschung*, 1, pp. 1-323 [trad. it. a cura di E. Filippini (1965) *Idee per una fenomenologia pura e per una filosofia fenomenologica. Libro primo. Introduzione generale alla fenomenologia pura*. Torino: Einaudi].

Irvine, J. T. and S. Gal. (2000) Language Ideology and Linguistic Differentiation, in P. Kroskrity (ed.) *Regimes of Language: Ideologies, Polities, and Identities*, Santa Fe, N.M.: School of American Research Press, pp. 35-83.

Kroeber, A. L. (1899a) 'Animal Tales of the Eskimo'. *Journal of American Folklore*, 44, pp. 17-23.

– (1899b) 'The Eskimo of Smith Sound'. *Bulletin of the American* Museum of Natural History, 12, pp. 265-327.

– (1899c) 'Tales of the Smith Sound Eskimo'. *Journal of American Folklore*, 12, pp. 166-82.

Labov, W. (1969) The Logic of Nonstandard English, in J. Alatis (ed.) *Georgetown Monographs on Language and Linguistics*, vol. 22, Washington, D.C.: Georgetown University Press, pp. 1-44.

Lee, D.S. (1979) *Native North American Music and Oral data. A Catalogue of Sound Recordings, 1893-1976*. Bloomington & London: Indiana University Press.

Liss, J. E. (1996) German Culture and German Science in the *Bildung* of Franz Boas, in G.W. Stocking Jr. (ed.) *Volksgeist as Method and Ethic. Essays on Boasian Ethnography and the German Anthropological Tradition*, Madison, WI: The Univeristy of Wisconsin Press, pp. 155-84.

McGee, W.J. (1897) Bureau of American Ethnology, in G.B. Goode (ed.) *The Smithsonian Institution 1846-1896. The History of the First Half Century*, City of Washington, pp. 367-96.

Morgan, M. (1994) The African-American Speech Community: Reality and Sociolinguistics, in M. Morgan (ed.) *Language & the Social Construction of Identity in Creole Situations*, Los Angeles: Center for Afro-American Studies, UCLA, pp. 121-48..

Philips, S.U. (1983) *The Invisible Culture: Communication in Classroom and Community on the Warm Springs Indian Reservation*. New York: Longman.
Sanjek, R. (ed.) (1990) *Fieldnotes: The Makings of Anthropology*. Ithaca: Cornell University Press.
Scheper-Hughes, N. (2001) 'Ishi's Brain, Ishi's Ashes: Anthropology and Genocide'. *Anthropology Today*, 17, pp. 12-8.
Steward, J.H. (1961) 'Obituary. Alfred Louis Kroeber. 1876-1960'. *American Anthropologist*, 63, pp. 1038-87.
Stocking, G.W.J. (ed.) (1974) *The Shaping of American Anthropology, 1883-1911: A Franz Boas Reader*. New York: Basic Books.
Wickwire, W. (2001) 'The Grizzly Gave Them the Song. James Teit and Franz Boas Interpret Twin Ritual in Aboriginal British Columbia, 1897-1920'. *American Indian Quarterly*, 25, pp. 431-52.
Wittgenstein, L. (1958) *Philosophische Untersuchungen/Philosophical Investigations*, trad. ing. di G.E.M. Anscombe, third edition, with English and German Indexes, The Macmillan Company [trad. it. di E. Filippini (1967) *Le ricerche logiche*, Torino: Einaudi].
Yates, M. (1982) 'Percy Grainer and the Impact of the Phonograph'. *Folk Music Journal*, 4.

La lingua parlata del teatro di narrazione
Analisi di tre monologhi (Paolini, Celestini, Enia)

Carmela Lucia
Università degli Studi di Salerno
Dipartimento di Letteratura, Arte e Spettacolo
carmenlucia_99@yahoo.it

Abstract

The subject of the present research is the analysis of the monologic speech in the "narrative theatre". The description of the "corpus" of theatre texts – enriched by the intertextual confrontation between the printed text and the performance, given by the audiovisual recording – is primarily oriented towards the analysis of lexicon and syntax, the aspects of enunciation, in particular repetition, the alternation of different types of deixis. This linguistic research, on the one hand, allowed the analysis of the way language has been changing in contemporary theatre under the influence of slang and dialects – Italian contaminated by the dialect of Veneto in Paolini, Roman in Celestini, regional Italian of Palermo in Enia – on the other hand, the study of interesting cases of interference between written and spoken language, from a certainly new and in some ways privileged point of view that is the monologue.

1 Premessa

Per la sua specifica fisionomia pragmatico-testuale, il testo teatrale si configura da sempre come il genere letterario sicuramente più vicino all'oralità, considerata l'esigenza di fedeltà mimetica al parlato, necessaria in funzione della resa performativa attoriale e, in generale, dell'*actio* drammatica[1]. La sua costituzionale complessità semiotica fa sì che questa specifica tipologia testuale si connoti come un insieme sinestesico di segni, dove si trovano a convivere, insieme con la scrittura propria del copione, le informazioni soprasegmentali, quali il ritmo, le pause e l'intonazione della voce, unitamente ad altri sistemi di segni, come, per esempio, i codici prossemici, cinesici o iconici, pertinenti alla *performance* scenica.

[1] Per avere un quadro di riferimento teorico e metodologico degli studi sul testo teatrale, analizzato dal punto di vista semiotico, si vedano in particolare Segre (1984: 3-15); Trifone (2000: 9-19).

Per questi aspetti, il testo teatrale si configura, di fatto, come un genere di confine, «diamesicamente ambiguo» (Trifone, 2000: 19), a metà tra i due poli delle due categorie esemplari dello *scritto-scritto* e del *parlato-parlato*, in quanto si presenta come divaricato tra il livello della *fabula agenda*, relativa alla pagina scritta, e l'oralità propria della sua attualizzazione scenica, pertinente al piano della *fabula acta*[2]. Ovviamente si tratta, in questo caso, di un'oralità «secondaria», per il processo di stilizzazione (o «filtraggio»), a cui inevitabilmente la pagina scritta sottopone il *parlato-recitato*, incluso nella più comprensiva categoria del parlato-simulato (D'Achille, 1990: 13).

Ebbene, sull'opportunità metodologica di analizzare la lingua "di scena" hanno già fatto luce importanti e numerosi studi di settore (Nencioni, 1983[1976]; Trifone, 1994, 2000; D'Achille, 1990; Stefanelli, 2006; Grignani, 2007), che hanno ribadito come il confronto tra il parlato programmato e il parlato autentico riveli quali spinte innovative e caratteristiche generali si possano coagulare in una tipologia testuale così specifica e, al contempo, ambigua come il copione teatrale[3].

In realtà, come verrà dimostrato nella presente ricerca, il testo teatrale può essere inteso come un fertilissimo *campo di tensioni* (per usare la nota metafora praghese), in cui si manifesta una dialettica di forze, che da un lato appaiono vincolate dal carattere di *permanenza*, proprio della scrittura – la *βεβαιότης*, di platonica ascendenza (Platone, 2000: 202) – dall'altro vanno a interferire con la struttura fortemente instabile e immediata (l'*evanescenza*) dell'esecuzione orale. Questo assunto è valido tanto più oggi, in quanto la possibilità di confrontare i

[2] Risale al noto studio di Nencioni (1983[1976]: 126) la formula esemplare del *parlato-scritto*, pertinente alla mimesi letteraria del registro orale della lingua, insieme al *parlato-recitato* proprio del testo teatrale. Per una discussione preliminare di questi aspetti rimandiamo alle pagine introduttive del nostro saggio (Lucia, 2008: 859-93).

[3] In particolare D'Achille (2001: 183), in un'importante indagine sulla lingua del teatro italiano dal dopoguerra a oggi, sostiene che «il confronto con il parlato autentico consente […] di cogliere sia le differenti percezioni del parlato che hanno avuto i singoli drammaturghi, operanti in momenti e situazioni diverse, sia le varie modalità con cui hanno cercato di fissarlo sulla pagina destinata alla scena. Non di rado, inoltre, i testi teatrali novecenteschi documentano forme, voci e costrutti che hanno attratto l'attenzione dei linguisti e dei grammatici solo successivamente; […] pertanto […] queste attestazioni scritte possono assumere – sul piano della storia linguistica – un significato non trascurabile».

due livelli testuali, mediante la collazione del testo scritto e della *performance* orale, registrata su supporti multimediali, consente di superare il limite sotteso all'analisi di un genere come quello teatrale, che, per la sua genesi scritta, è stato a lungo escluso (o comunque tenuto ai margini) dalle analisi sul parlato, nonostante sia ormai un dato ormai consolidato per le ricerche di questo settore che: «la lingua in scena costituisca un interessante osservatorio "sui generis" delle relazioni tra pragmatica, testualità e grammatica» (Trifone, 1995: 194).

2 Il teatro di narrazione: sull'opportunità metodologica di indagarne la lingua (limiti e vantaggi della ricerca)

Oggetto della ricerca, che si presenta in questa sede, è l'analisi della lingua del monologo del "teatro di narrazione" – un complesso insieme di fenomeni teatrali, diffusisi in Italia intorno alla fine degli anni Ottanta[4] – che si fa coincidere con la riscoperta di una drammaturgia parlata, con la valorizzazione della dimensione orale della lingua e il recupero della "letteratura orale" di tradizione popolare.
Caratteristica principale dei cosiddetti "monologanti" – come Ascanio Celestini, Marco Baliani, Marco Paolini, Laura Curino, Davide Enia, Gabriele Vacis, Alessandro Bergonzoni, Lella Costa, Moni Ovadia, Giancarlo Pirrotta e altri – è quella di essere accomunati da una generica matrice culturale, ovvero da una prassi recitativa che, nonostante le diversità idiolettali e artistiche di ciascuno di essi, si orienta verso un'unica soluzione linguistica, vale a dire verso l'adozione di un linguaggio piano ed essenziale, connotato da forti spinte verso l'oralità, riconoscibile soprattutto per l'adibizione di registri e varietà linguistiche medio-basse, talora marcatamente regionali, se non addirittura contaminate dai dialetti di appartenenza.

[4] Secondo Guccini (2005: 11): «il "teatro narrazione" è stata una delle etichette più in voga all'inizio degli anni Novanta, e, da allora, anziché eclissarsi o ripetersi perdendo lo smalto della novità, è piuttosto diventato una possibilità diffusamente praticata e reinterpretata dal mondo teatrale, un *quasi genere*, che, in un ideale palinsesto delle possibilità scenico-performative, affianca il "teatro ragazzi", il "teatro danza", il "teatro immagine", la "sperimentazione", il "teatro di prosa" e il "teatro nel sociale" ».

In sostanza, per quanto sia classificabile pur sempre nella categoria del *parlato-recitando* (Nencioni, 1983[1976]: 127) o anche dell'*oralità spettacolare* (Trifone, 1994: 18), la "lingua parlata" dal *teatro di narrazione* – diversamente dagli altri generi teatrali, in virtù della sua particolare genesi compositiva, si mostra affrancata quasi del tutto dal vincolo tradizionale con lo scritto, in quanto la composizione testuale (la cui stesura è successiva a un'intensa fase di elaborazione orale) assume di fatto un carattere "consuntivo"[5]. In questo caso, il copione non si identifica più con un "modello" che l'attore deve seguire, quanto piuttosto viene a configurarsi come una traccia – suscettibile di continue riformulazioni e cambiamenti – che, di volta in volta, può mutare di segno. Di conseguenza, la scrittura, vale a dire il momento formale della stesura testuale del monologo, non sempre precede la *performance* orale, ma, anzi, diventa rispetto ad essa un canale secondario. In sostanza, la pagina scritta, o livello della *fabula agenda*, viene ad assumere così un ruolo vicario rispetto alla recitazione, perché si alimenta di invenzioni estemporanee, proprie del racconto orale, e di nuove riformulazioni, che nascono nelle diverse *performances* degli attori[6].

Pertanto, è ovvio che si inverte il tradizionale processo, che dalla pagina scritta porta alla *performance* attoriale, rendendo ancora più stretto il varco tra la lingua scritta e il parlato.

Ora, tutti questi aspetti, che investono la particolare genesi compositiva dei monologhi del teatro di narrazione, provocano conseguenze importanti, da valutare attentamente ai fini di una ricerca, che si pone come obiettivo l'analisi del parlato: in quanto la rielaborazione continua del testo, alla luce della pratica viva dello spettacolo, fa sì che il rapporto tra scrittura e oralità non si possa più

[5] I termini di drammaturgia *consuntiva* e *preventiva* sono stati coniati da Ferrone (1992: 97), e poi riutilizzati da Ariani and Taffon (2001: 285), per indicare un tipo di teatro (comune a tanti gruppi sperimentali degli ultimi decenni) dove la composizione testuale, nel primo caso, segue quanto già è stato articolato nella messa in scena; nel secondo caso, per drammaturgia *preventiva* si intende invece una modalità compositiva tradizionale, lì dove il copione precede la recitazione e gli allestimenti scenici.

[6] Molti monologanti provano più volte i loro spettacoli in circoli privati, piccoli centri, laboratori di sperimentazione teatrale, per testare la ricaduta dei racconti sulla platea; perciò il pubblico diventa spesso un pre-selettore, che incide anche sulle scelte linguistiche adottate nei racconti. Per questi aspetti, si veda Guccini (2005: 12-26).

concludere nel rigido monolitismo delle tradizionali e pacifiche bipartizioni (scritto *versus* parlato), soprattutto se si considerano le importanti interferenze e le molteplici dinamiche, che si intrecciano nelle complesse traduzioni intersemiotiche, nel passaggio dalla pagina alla scena e viceversa.
Una volta considerati questi complessi fattori che interagiscono nella genesi del monologo del teatro di narrazione, bisogna valutare alcuni assunti di base, che giustificano la scelta di questa particolare tipologia testuale come oggetto d'analisi della presente ricerca:

- innanzi tutto il monologo nasce da una scrittura oralizzante, in quanto il *performer* (l'autore-attore che interpreta il monologo), nell'atto compositivo e genetico del testo, che passa attraverso numerose prove e attività di laboratorio, condivide con lo spettatore la messa a punto del testo, facendola coincidere con l'atto stesso della narrazione.

> All'assenza di un testo "preventivo", che fissi una volta per tutte i contenuti verbali del racconto, corrisponde, in corso d'opera, la presenza d'una molteplicità di testi, che alimenta e indirizza l'azione orale (Guccini, 2005: 23).

In generale, soltanto dopo aver provato più volte e dopo aver portato sulle scene il monologo, l'autore-attore si limita a trascrivere gli esiti della composizione orale sulla pagina e, molto spesso, non usa trascrivere il copione, che rimane quasi sempre provvisorio.

- Di conseguenza, nella traduzione o trasmutazione diamesica dei segni verbali in "*gramma*"[7], in questo caso è «l'oralità-che-si fa testo» (Guccini, 2005: 24)[8] e non viceversa; dal momento che il monologo

[7] Cfr. Jakobson (1966: 56-64). Nel processo di «traduzione dal verbale orale al verbale scritto e viceversa si parla di "traduzione endolinguale diamesica"» (Petrilli, 1999: 422).
[8] Scrive giustamente Guccini (2005: 24-5): «Se la "scrittura oralizzata" è una tecnica ben nota ai romanzieri e agli autori drammatici, ciò che qui definiamo come oralità-che-si-fa-testo è possibilità esclusiva di chi – narratore, comico, improvvisatore – condivide con lo spettatore la messa a punto del testo, facendola coincidere con l'atto della narrazione. In questa fase, il *performer* sedimenta le indicazioni che emergono dal pubblico, i propri impulsi e le innovazioni estemporanee che

appare svincolato dalle tradizionali convenzioni normalizzatrici della scrittura, in quanto non è più concepito endofasicamente sulla pagina, in vista della recitazione, ma nasce già in una dimensione orale, che verrà poi solo successivamente fissata in quello che sarà il copione (D'Achille, 2001: 185).

Soprattutto quando si confronta il copione con la registrazione audiovisiva si verifica spesso che il testo di arrivo, effettivamente recitato, non riproduca fedelmente la lingua del testo originale, dal momento che il monologo non si identifica più, per l'attore, come un modello da rispettare, quanto, piuttosto, come «una variante documentata e adeguata alle esigenze della lettura» (Guccini, 2005: 22).

Quindi, diversamente dalla prassi canonica, valida per la drammaturgia "consuntiva", dove il sistema dell'oralità scritta condiziona e vincola la *performance* orale, rendendo il parlato più sorvegliato e normativo, nella lingua del teatro di narrazione il coefficiente d'oralità appare sicuramente più alto e marcato, per l'espressività del parlato e l'ispessimento della filigrana dialettale, che incidono notevolmente in un testo orale, dove è ammesso di fatto più margine di improvvisazione.

3 Il teatro di narrazione nella drammaturgia contemporanea

Dopo aver evidenziato queste premesse teoriche, necessarie a inquadrare le implicazioni metodologiche della presente ricerca, è opportuno delimitare brevemente un profilo descrittivo del teatro di narrazione, nel quadro della drammaturgia contemporanea.
Innanzitutto, occorre chiarire che, diversamente dal teatro borghese, la lingua dei *nuovi performer epici* [9] – come Fo, Celestini, Paolini, Baliani, Enia – esibisce varietà linguistiche medie o medio-basse, perché porta in scena realtà sociolinguistiche e livelli di cultura diversi

punteggiano lo spettacolo: accelerazioni, sospensioni, associazioni, accentuazioni di senso, nuove immagini, raccordi e dialoghi improvvisati».

[9] Si deve a Meldolesi and Guccini (2004: 3-4) questa definizione, che inquadra un gruppo d'autori, accomunati da una generica matrice culturale, nonché da una prassi scrittoria, che si orienta verso un'unica soluzione linguistica, nonostante le diversità idiolettali e artistiche di ciascuno di essi.

dall'italiano letterario, lontani da una visione "grammaticocentrica" della lingua, fondata sulla dominanza dei valori della scrittura letteraria e della sua tradizione formale.
Dopo la preziosa lezione di Eduardo De Filippo e Dario Fo, «eredi e interpreti del retroterra comico regionale di appartenenza» (Stefanelli, 1987: 247), ma anche di attori, interpreti di una lunga tradizione «mattatoriale» – da Pratolini a Proietti e Franca Valeri (Stefanelli, 2006: 31) – questa nuova generazione di autori-attori inizia a guardare la specificità della lingua di appartenenza, impastata di *colore locale* e, in particolare, a recuperare l'oralità dialettale, che si connota, per molti di loro, come una lingua allusiva, evocativa, simbolica.
Modello per antonomasia del genere narrativo affabulatorio è sicuramente Dario Fo e la sua "scrittura d'attore", ovvero un particolare tipo di drammaturgia, che si alimenta della tradizione specifica e dell'esperienza diretta della ribalta. Dalla sua lezione i monologanti della generazione degli anni Ottanta ereditano l'attenzione per l'attualità politica e sociale, nonché l'attività di recupero e ricerca di diversi filoni del teatro popolare. Ciò nonostante, diversamente dalle scelte eversive dell'istrionico *performer*, i nuovi autori-attori optano per l'adozione di un registro narrativo, che miri sempre alla massima comprensibilità, in cui l'uso dell'oralità dialettale non risulta così contaminato da «scelte violentemente antinormative» (D'Achille, 2001: 186), rispetto allo standard.
Com'è noto, nella funambolica lingua e nel *pastiche* del magma pluridialettale dei testi di Dario Fo, si mescidano «elementi volgari vivi del suo lombardo occidentale e di altri dialetti dell'area settentrionale padana lombardo-veneta, talora friuliano», con «materiali di riporto della tradizione giullaresca» (Folena, 1991: 120-21), mentre la lingua del teatro di narrazione appare contrassegnata da scelte linguistiche sicuramente meno eversive e devianti rispetto alla norma.
In generale, alla vocazione plurilingue del teatro italiano sono da sempre appartenuti l'uso del dialetto, come lo sfruttamento drammaturgico delle varie *koinaí* regionali, o la contaminazione espressionistica, con le sue violazioni antinormative, che dal maccheronico arriva a certi Scapigliati fino a Gadda, la cosiddetta linea "Folengo-Gadda", di continiana memoria, che punta sull'espressionismo verbale di una lingua artificiale, contaminata da lombardismi, arcaismi, latinismi, lessico basso di origine popolare. E

del resto questo aspetto si riflette ancora oggi nella drammaturgia contemporanea, in cui si registra sicuramente una nuova vitalità dell'oralità dialettale: si pensi, per esempio, alla scuola partenopea, con i suoi principali esponenti: Cappuccio, Moscato, Ruccello, Santanelli; alla Sicilia di Scimone, Pirrotta, Cuticchio, Enia; alla lingua toscana di Ugo Chiti; al romano di Manfridi, Marino, Tirabassi.
Per questi fattori, l'attenzione dei linguisti si è spostata sull'analisi dell'oralità dialettale a teatro (lo testimoniano le importanti ricerche di: D'Achille, 2001; Giovanardi, 2002, 2006; Stefanelli, 2006, 2007), dal momento che come sostiene giustamente Stefanelli (2006: 18):

> i drammaturghi novecenteschi hanno gradualmente ma progressivamente maturato una coscienza linguistica nazionale che consiste nel possesso di un italiano parlato libero da rigidezze scolastiche di ascendenza letteraria ma nello stesso tempo consapevole della propria memoria dialettale e aperto agli influssi delle altre lingue.

Proprio in virtù di ciò, è importante quindi orientare l'attenzione verso l'attività dei monologanti, per il progressivo rivolgersi della loro drammaturgia alla specificità della lingua locale e agli italiani regionali, diventati ormai uno dei motori principali del successo e dell'esplosione vitalistica di questo tipo di teatro.

4 Metodologia di ricerca e descrizione del corpus analizzato

Alla luce di tali considerazioni prodromiche, è necessario, a questo punto, illustrare la metodologia della ricerca e i parametri di analisi, utilizzati per il corpus dei testi selezionati.
Come si è anticipato nelle pagine precedenti, è importante considerare questi punti fondamentali:

per il teatro di narrazione, le condizioni di produzione del testo sono diverse rispetto alla prassi compositiva del copione della drammaturgia tradizionale. Spesso non esiste un testo *preventivo*, dal momento che l'attore narrante – protagonista, testimone e cronista della storia – segue non un copione, quanto piuttosto un canovaccio, che solo dopo numerose prove e modifiche approda a una stesura definitiva. Perciò i monologhi si qualificano, anche a una prima lettura sommaria, per la loro spiccata vocazione orale, che coincide con

l'avvicinamento al parlato spontaneo e con un forte ispessimento della filigrana dialettale;

un secondo dato incontrovertibile, che interessa valutare in questa sede, è che l'elemento del racconto epico, vale a dire la dimensione orale della lingua enfatizzata dalle *performances* dei monologanti, si identifica di fatto con una nuova modalità di comunicazione, in quanto il narratore, sostituendosi al giornalista e alla sua funzione testimoniale, restituisce al pubblico argomenti di attualità politica, recuperati spesso dalla storia recente e rielaborati sul filo di una prassi comunicativa, che coincide con la riscoperta di una drammaturgia "parlata, più che agita". Per questo il *performer* monologante, unendo testi di tipo *fictional* e *non fictional* (vale a dire, storie inventate, con documenti storici, recuperati dalla memoria collettiva), comunica senza filtri, senza maschere di altre *dramatis personae*, direttamente con lo spettatore, parlando del proprio vissuto in prima persona, fino a sostituire «al principio della *messa in scena* […] quello della *messa in contatto*» (Guccini, 2005: 12).

4.1 Il corpus dei monologhi

Dopo aver considerato questi aspetti, è necessario ora illustrare quali testi compongono il corpus dell'analisi. Sono stati scelti i monologhi di tre autori, che possiamo considerare esemplari:

- Marco Paolini e Gabriele Vacis (2008) *Vajont, 9 Ottobre '63. Orazione civile*. Torino: Einaudi;
- Ascanio Celestini (2006) *Scemo di guerra. Il diario*. Torino: Einaudi;
- Davide Enia (2005) *Italia-Brasile 3 a 2*. Milano: Ubulibri.

I testi, corredati della registrazione audiovisiva (tranne nel caso di Enia), riflettono con la loro lingua una sorta di *geografia dei dialetti in scena*: nel caso di Paolini la lingua insiste su un sostrato dialettale veneto; per Celestini il dialetto d'appartenenza è il romano; mentre per Enia il palermitano.
In particolare, per Ascanio Celestini e Davide Enia la scelta di usare un registro medio o basso, se non addirittura gergale, coincide con la ricerca della memoria storica, intrecciata all'idioletto familiare.

Diversamente, per Marco Paolini – che usa «un italiano regionale veneto di ampia comprensibilità, funzionale all'esigenza comunicativa dominante sulla scena» (Stefanelli, 2006: 32) – i temi tratti da documenti storici spesso trovano alimento nelle attività pedagogiche e formative dei gruppi sperimentali, a cui l'autore appartiene[10].

4.2 Metodologia d'analisi

Sul piano metodologico, si è scelto di selezionare e di discutere gli aspetti della lingua, che ricorrono nei monologhi con maggior frequenza, sino a certificare ─ con le dovute modularità e nel rispetto delle diversità di ispirazione e artistiche dei singoli autori ─ tratti comuni, che si caratterizzano per la spiccata vocazione orale e l'avvicinamento della lingua del teatro al parlato spontaneo.
Inoltre, la possibilità di confrontare testo scritto e *performance* orale, mediante la metodologia della collazione tra le due versioni, vale a dire il confronto sistematico tra copione e trascrizione linguistica ─ resa possibile grazie alle registrazioni dei media audiovisivi (in VHS e DVD) e già sperimentata nella nostra precedente ricerca (Lucia, 2008: 859-93) ─ ha consentito di allargare e potenziare lo spettro d'analisi dato dalla rappresentatività del parlato teatrale.
È superfluo infine sottolineare che presenteremo, in questa sede, solo un numero limitato di campioni, con una selezione di esempi indicativi, mutuati da sistematiche schedature dei testi a stampa, confrontati con la registrazione su supporto multimediale.

5 Analisi dei dati

5.1 Marco Paolini: la narrazione "civile" tra epos e storia

Marco Paolini (Belluno, 1956) è noto al grande pubblico soprattutto grazie al successo mediatico del monologo *Vajont, 9 ottobre '63*.

[10] Gli stessi autori-attori dimostrano inoltre di avere una forte consapevolezza metalinguistica in merito, in quanto, mentre sperimentano la forza enunciativa dei monologhi, ispirati sempre ad un principio di verità, riflettono, nelle loro argomentazioni epilinguistiche, sul carattere diretto del parlato, basso e socialmente marcato delle scelte di registro (Guccini, 2005: 22-43).

Orazione civile (2008), che dalla diretta in prima serata su RaiDue (del 9 ottobre 1997)

influisce coi suoi tre milioni e mezzo di telespettatori anche sugli sviluppi del "teatro narrazione" accelerandone la trasformazione in un "quasi genere" di riferimento (Guccini, 2005: 135).

Il monologo *Vajont*, in cui si intrecciano senso civile, oralità e drammaturgia, è un'orazione civile, che rievoca le vicende storiche e politiche della tragedia del crollo della diga del Vajont, con la ricostruzione, dal progetto al processo, della sua complessa e oscura storia. Il testo fa rivivere, in presa diretta, i momenti più intensi di una tragedia nazionale, raccontata da un "narratore civile", come Paolini, che unisce la recitazione d'impronta epica, rivolta al pubblico e animata dalle sue straordinarie doti mimiche e recitative, all'esattezza della testimonianza storica, della memoria e della denuncia sociale[11].
Presentiamo qui una campionatura degli esiti morfosintattici parlati più interessanti, tratti da *Vajont*, con lo spoglio linguistico, che emerge dal sistematico confronto tra testo a stampa e testo recitato (registrato in DVD, della durata di 157').
Dalla trascrizione dei dati linguistici, si evidenzia che, nel passaggio dalla scrittura alla *performance* orale: per la morfologia verbale si verifica un aumento considerevole delle forme composte dei verbi del tipo *vien giù*; *vien giù sassi*; numerose apocopi postconsonantiche, come documentano i seguenti esempi: *son tre ore che è là*; *fatta far questa fatica*; *non son stati distrutti*; *son preoccupato*; *vien preso un provvedimento*; *son sicuro.* Inoltre, aumentano i casi in cui è evidente l'uso dell'articolo determinativo davanti al nome proprio, come nell'esempio seguente: *la Tina Merlin.*

[11] Nel fondamentale saggio di Guccini sul teatro di narrazione, è ricostruito il complesso processo genetico che sta a monte dei monologhi di Paolini, monologhi che, come nel caso di *Vajont*, nascono da un intenso lavoro di studio e riscrittura di documenti storici o romanzi. Intrecciando le funzioni del narratore a quelle del testimone e dell'autore orale, Paolini rinuncia alla rete delle collaborazioni artistiche, attoriali e registiche, che caratterizzavano le sue prime esperienze teatrali, per costruire il monologo *Vajont* come un *work in progress*, in una serie di incontri in circoli privati, piccoli teatri, «in quanto esperienza "non progettata" e via via costruita nella successione degli incontri» (Guccini, 2005: 139).

Si registra una più alta frequenza anche di alcune occorrenze dell'uso di *mica*, con omissione dell'avverbio negativo: *si può mica spender*; *non avrò mica già litigato col mio papà alle sette di mattina*; *non avevamo mica parenti noi a Longarone*.
Per le forme di attestazione pan-italiana, si evidenzia l'uso aferetico di *questo*, come nell'esempio, *'ste stazioni*; o anche, sempre per i moduli della deissi, l'uso enfatico dei dimostrativi: *questa diga qua*; *questi altri qui.*
Un elemento importante tipico della ridondanza del parlato è sicuramente l'iterazione dell'aggettivo, come evidenzia uno dei numerosi esempi rubricati: *bello, bello, bello*. Sul piano sintattico si verifica un considerevole aumento delle dislocazioni: *ma quelli a casa / c'è qualcuno che se li è dimenticati*; *io / le stazioni le avevo imparate a memoria.* Alcuni esempi di costrutti con *nominativus pendens*: *la commissione / io ci ho messo l'aggettivo / ma ovviamente il ministro il timbro.* Sempre nel passaggio dalla pagina alla registrazione in DVD, aumentano le allocuzioni all'ascoltatore (*ricordate?*; *dicevo*; *sai*), che testimoniano quasi un'apertura del discorso verso il pubblico, continuamente stimolato all'ascolto, quasi convocato sulla scena, per riannodare insieme all'attore i fili del racconto.
Si registra infine un notevole aumento delle interiezioni e, in generale, dei segnali discorsivi, nel passaggio dalla pagina scritta all'oralità dell'attualizzazione scenica, fenomeno già rilevato nelle precedenti ricerche (Stefanelli, 2007: 343; Lucia, 2008: 885). Nella *performance* attoriale risulta molo alta la frequenza delle interiezioni e delle pause vocalizzate, forme semanticamente vuote, usate per modulare la voce, che sfrutta a pieno il loro valore ritmico e musicale, come per "saturare" il testo orale, aprendo e chiudendo le unità prosodiche.

5.2 Ascanio Celestini: l'oralità popolare delle borgate

I monologhi di Ascanio Celestini (Roma, 1972) si ispirano all'arcaica semplicità del racconto orale, espressione del *milieu* romano d'appartenenza, che si intreccia con i frammenti di una memoria personale e familiare, conclusa nel ricordo di frantumi vocali, recuperati dalla memoria del suo speciale *lessico famigliare*.

La tacitazione delle componenti esornative della lingua coincide con la ricerca di un'espressività parlata, che si basa sull'elementarità discorsiva dell'oralità popolar-regionale e soprattutto sul recupero dei dialettalismi gergali. Diversamente da Paolini, che sceglie un registro linguistico di tono medio, la lingua di Celestini, caratterizzata da una spiccata velocità di eloquio, appare più marcata dal punto di vista diatopico, anche per la presenza di figure evocate sulla scena, il popolo delle borgate, vere e proprie *dramatis personae*, espressione di un gruppo omogeneo dal punto di vista diastratico. La forte specificità diatopica del materiale verbale – un «italiano popolare (o italiano dei semicolti) di stampo romanesco» (Stefanelli, 2007: 340) – è espressione del microcosmo di riferimento rappresentato, la popolazione del Quadraro, rastrellato nel '44 durante la Seconda Guerra mondiale.

Come si è anticipato, il testo scelto, confrontato con la registrazione in DVD (della durata di 135') è *Scemo di guerra. Il diario 2006-1944* (2006): sul piano strettamente compositivo appare costruito sulla base di verosimili coordinate antropologiche, tipiche della mentalità contadina e delle borgate, che si riflettono nell'intento mimetico e nella tensione realistica del racconto.

Nella partitura oralizzante del monologo il livello espressivo comporta, per l'addensarsi di moduli popolareschi, un ispessimento dell'oralità dialettale e un recupero di forme idiomatiche, sfruttate soprattutto per il loro alto valore mimetico, in quanto tipiche di un registro diafasicamente basso. Una notevole frequenza registrano gli schemi anaforici – veri e propri *patterns* tipici della narrazione orale – che dominano le procedure compositive del racconto e lo stile "formulaico", proprio dei racconti orali (Lavinio, 1990: 133-56).

In generale, l'andamento discorsivo si presenta privo di forti nessi subordinativi, per il reiterarsi delle congiunzioni semplici, come per esempio la *e*, o il *che* usato in funzione di un generico giuntore coordinativo, come è evidente nell'esempio seguente: *Sisto Quaranta dice che il Quadraro è un quartiere isolato, per questo che i tedeschi hanno potuto fare il rastrellamento… che forse è stato il più grande che sia stato fatto a Roma* (si cita dal testo a stampa *Memorie*, p. 47).

Sul piano della *performance* orale si riscontra una notevole velocità di esecuzione nella recitazione, con effetti di sincopato, ma anche essenziali partiture ritmiche, in un quadro per certi versi anadiplosico, con ripetizione delle stesse parole anche a breve distanza.

Il procedimento dominante del monologo è sicuramente paratattico, con prevalenza del modulo asindetico e segmentazione del periodo in unità irrelate, prive di più articolate giunture subordinative: attraverso schemi anaforici il racconto appare effettivamente costruito con legami semplici, dati dall'iterazione delle stesse parole. Così nel segmento testuale: *Ci sta gente come Sisto, che fa l'elettricista, che al momento del rastrellamento stava dormendo e ha pure provato a spacciarsi per malato… ma se lo sono portato via lo stesso. Ci stanno operai e artigiani, qualche panettiere e qualche pasticcere che alla fine della guerra, se tornerà vivo a casa, al massimo potrà raccontare alla moglie e ai figli quello che gli è successo.*
Ma non ci stanno i poeti e gli scrittori.
Non ci stanno gli intellettuali che scrivono sui giornali, che pubblicano libri.
Questo tipo di legame semplice configura un andamento discorsivo, in cui il rapporto diretto che lega il parlante con gli eventi si esprime con un procedere a scatti, in uno sviluppo orizzontale, fitto di pause e di continue approssimazioni. Si veda il seguente esempio: *E proprio per questo non è possibile che quelle sono papere vere. 'Na papera vera fatta di carne viva di papera si brucia a stare su quell'acqua incendiata. Una papera vera scappa dall'acqua che scotta.*
Come si è detto, sul piano lessicale la lingua di Celestini appare più marcata sul piano diatopico, per l'adibizione di un impasto di stampo romanesco-gergale:

da qui l'uso di forme come: *mi padre*; *mi zio*; *mi nonno*; *mi moglie*;
l'adozione del possessivo posposto nei sintagmi: *alla madre sua*; *un parente suo*; *un parente mio*; *il marito mio*; *il negozio suo*; *tutti i figli miei*;
per la morfologia nominale, l'uso della *e* epitetica, come in *tranve*;
forme morfologiche apocopate: *dei spaghetti*; *i strilli*; *sull'ipotetici soci*; *dell'antichi romani*; *'na porta*; *'na finestra.*
per la morfosintassi, l'uso di *manco*, in funzione di forma negativa: *non sapeva manco bene*; *non era manco tanto sicuro*;
o infine, l'uso della *a* allocutiva, in espressioni come: *'a Sisto*[12].

[12] Per questi aspetti è importante il confronto con il fondamentale studio sul "romanesco nel teatro contemporaneo" di Giovanardi (2006: 107-28).

Inoltre, largamente rappresentati sono i morfemi a forte connotazione locale, di cui si presenta una breve campionatura, con idiotismi e locuzioni popolari: *la cagiara*; *si capava una mela*; *stanno a fa' 'e lumacate.*

Sul piano della morfologia verbale si evidenziano numerose apocopi: *vede'*; *chiacchiera'*; *arriva'*; *a rompe'*; *non devo ammazza'*; al congiuntivo imperfetto si trova il morfema *–ssi –sse* nella terza persona singolare: *il barbiere quando lo vide glie parve che l'acqua stasse ferma*; si registra inoltre l'uso dell'indicativo per il congiuntivo: *i tedeschi ancora non se ne erano annati / de modo che chiunque passava.*

In generale, i dimostrativi si presentano quasi sempre nelle forme aferetiche, un tratto dell'italiano dell'uso medio (Sabatini, 1985: 158), che rende evidente il tentativo di trasferire, anche sulla pagina, alcuni tra i fenomeni di "allegro", spesso usati nel parlato spontaneo, non pianificato: *'sta cosa*;*'sto cane*; *'sto tetto*; *'sta finestra*; *'sta bestia*; *'sto maiale*; *'sti parenti*; *'sto cinema*; *'sto barbiere.*

Sempre sul piano morfosintattico, notevole è la frequenza dei deittici usati con valore intensivo; i dimostrativi *quello*, o *questo* sono spesso accompagnati da altri segni indicali, come dimostra l'esempio che segue: *ma quello / mi padre / diceva che era una cosa normale*; *quello / mi nonno disse che lui*; *il parente / questo qui di Frascati.*

Analogamente si riscontra l'uso intensivo delle particelle pronominali, come il *ci* ridondante con il verbo *avere*: *mi padre / c'aveva otto anni.*

Per il livello sintattico, si registra un'alta frequenza di forme d'ingresso o connettivi testuali, come *allora*, *comunque*, *ma*, *e*, *o*, iniziali di discorso o di enunciato. Si documenta qui l'esempio di *e* enfatica introduttiva di principale, posta in attacco di frase: *e quest'uomo è morto che manco un figlio gli era lasciato alla moglie*; *e ha aperto una fioreria / che lui c'ha lavorato / campandoci dentro / fino al 1972.*

O anche l'uso di nessi istituiti dal *che* indeclinato e polivalente, con funzione sincretica, sostitutiva di una congiunzione con valore consecutivo, come appare evidente nel segmento qui riportato: *gente povera / che se uno muore da un giorno all'altra / i familiari non c'hanno il vestito da morto*; o con riduzione allomorfica *che/ in cui: e ha aperto una fioreria / che lui c'ha lavorato / campandoci dentro / fino al 1972 raccontava dei suoi nonni / per esempio /che mai s'era parlato dei nonni.*

Fenomeni strettamente connessi con la scarsa pianificazione dell'atto comunicativo orale sono, inoltre, i mutamenti di progetto e le dislocazioni, le segmentazioni frasali e i vari moduli della tematizzazione, con concordanze a senso e costrutti anacolutici: *se tieni aperti l'occhi / ti capiscono subito che sei vivo eh / questa storia / è la storia che ho raccontato tutta la vita / era / era proprio la sua / la sua storia insomma / quella che / quella che se dicevi*; *difatti / a mi padre nella memoria / questo tedesco / glie parve sempre una strana razza d'incrocio*; *le mani belle / al barbiere / gli erano venute / non perché tagliava barba e capelli / ma perché tagliava la carne.*

Si registra inoltre un'alta frequenza delle reduplicazioni e delle ripetizioni, che vanno intese sia come marche tipiche di un'oralità non pianificata, sia come modalizzatori, espressione dello stile "formulaico", tipico dei racconti orali e delle fiabe, come per esempio le più varie forme dell'intensificazione espressiva, realizzate con raddoppiamenti e ripetizioni inglobate in un tipo di sintassi esclamativa ed ellittica. Si evidenziano infatti numerose occorrenze di reduplicazione dell'aggettivo: *un soldato giovane giovane*; *l'uomo magro magro / secco secco.*

Le forme della *geminatio* rappresentano registri vocali tipici del parlato informale, lì dove la ripetizione di una stessa parola (a contatto, o a distanza) è intesa come manifestazione dell'*amplificatio* emozionale, con cui il parlante tende a focalizzare l'attenzione di chi ascolta su un'informazione. Oltre i raddoppiamenti intensivi, schema sonoro tipico dell'oralità, altri elementi tipici della ridondanza verbale sono le intensificazioni con *tricolon*, come il seguente: *la neve era rimasta fredda / ghiacciata / congelata.*

Un aspetto importante del racconto sono inoltre i toponimi, i nomi-parlanti e i *nicknames*, emanazioni verbali di un *humus* culturale, in cui la comunicazione appare veicolata dalle varie forme dell'oralità, come pure le locuzioni proverbiali, i modi di dire, che giocano sulla memoria del ritmo, nella scansione binaria o ternaria della loro formulazione, con stereotipi cristallizzati nella dimensione orale della lingua. Qualche esempio: *a Roma / come si dice / c'hai le mani di ricotta*; *e per fortuna che non s'è messo a «fare il bulletto» sennò rischiava di fare una brutta fine* (con trascrizione del raddoppiamento fonosintattico in *sennò*); per l'onomastica: *er sor Giulio*; o infine i moduli dell'alterazione sostantivale, con un uso ironico, o epitetico-

ostensivo dei diminutivi: *Gaetano Celestini, ma tutti lo chiamavano Nino.*
Unitamente ai modi di dire, simbolo di una *clôture* verbale, in cui circola un ipertrofico flusso di locuzioni stereotipe ed elementi formulaici, vanno considerati anche gli alterati a connotazione spregiativa, o gli epiteti, e l'insulto reciproco (*flyting*), comuni a tutte le società a cultura orale, che svolgono il ruolo di evidenziatori pragmatici e intensificatori della semantica dell'enunciato (così come testimonia l'esempio: *lo stesso ragazzino paraculo*).
Sul piano dell'enunciazione, si registra un'alta frequenza di moduli demarcativi d'enunciazione o interrogative, del tipo: *come no?*; *capito?*; *Che posso di'?*; *Figuriamoci*; *nun me ricordo*; o formule con funzione demarcativa, come per esempio: *vabbè*; *insomma.*
Il monologo, spesso percorso dal metronimico affioramento del *mo'* o il ricorso a *ecco*, è scandito in ampie sequenze esclamative regolate quasi ritmicamente dagli scatti interiettivi, come il *no*, usato in funzione di connettivo testuale e altri segnali metadialogici, uniti alle forme fàtiche, legate all'enfasi, come per esempio le interiezioni primarie, che aumentano nel passaggio dal testo scritto all'esecuzione orale, così come le pause di esitazione, le formule di ripetizione e di ripresa.
Come in altri monologhi, si distingue un *tic* linguistico, tipico di Celestini – già rilevato in precedenti ricerche (Stefanelli, 2006: 35) e (Lucia, 2008: 880) – vale a dire l'uso ripetuto dell'introduttore generico e desemantizzato *dice*. Assunto come un indice della bassa sinonimia, propria del parlato non pianificato, qui va considerato come uno stilema tipico della replica, poiché viene usato per segnalare il passaggio dal testo autoriale al discorso diretto simulato, come evidenziano questi esempi: *e poi dice che / l'americano si girò*; *e dice che erano trenta / quaranta soldati sbandati*; *«E chi'o pagava, er tranve», dice* (p. 93)[13].

[13] Nel vasto repertorio delle interiezioni, dei demarcativi, dei segnali di presa di posizione enfatica e, in genere, dei segnali discorsivi, l'iperonimo *dice* va qui inteso come un segnale rafforzativo dell'enunciazione e verbo introduttore neutro, che svolge la funzione di segnalare il passaggio delle voci, sia nelle forme del discorso diretto, sia indiretto.

5.3 Davide Enia: il ritmo arcaico del "cunto" nella scrittura oralizzata

Il giovane attore palermitano Davide Enia (Palermo, 1974) eredita dalle antiche affabulazioni popolari dei *cuntisti* siciliani la tecnica del *cuntu*, una narrazione di un fatto notevole, che il narratore ricuce sul filo della memoria autobiografica, con le tecniche narrative degli antichi aedi e dei cantastorie.
In effetti, i monologhi di Enia riproducono, nella dimensione ritmica della recitazione, il saliscendi vocalico proprio di questo tipo di racconto, secondo il linguaggio popolare e rituale del "cuntista" (Lezza, 2005: 7-13): nello specifico, il testo analizzato, *Italia-Brasile 3 a 2* (2005), rievoca l'epopea sportiva dell'Italia contro il Brasile, durante il mondiale di calcio del 1982, rivissuto attraverso la lente deformante dell'attore narrante. Rispetto alle scelte linguistiche di altri monologanti, che si attestano su opzioni di tono medio – come per esempio Paolini, o anche Marco Baliani – Enia sfrutta, si può dire, *ad libitum* l'espressività e la musicalità del dialetto siciliano[14].
Il registro narrativo che domina il monologo è fondamentalmente comico: l'attore racconta in prima persona la versione *d'en bas* di una cronaca sportiva della memorabile partita di calcio, vissuta da una famiglia palermitana, che consuma un'«orgia bacchico-dionisiaca davanti al teleschermo» (*'u moloch Sony Black Trinitron)* (*idem*, p. 35), narrata in una *climax* ascendente di *pathos*, dal *gòlle* annullato, fino alla vittoria insperata.
Il testo analizzato (senza il supporto della registrazione audiovisiva) si struttura come un discorso indiretto libero, con inserti e citazioni di discorso diretto, che instaurano una sottile trama polifonica, con un insieme di ibridi tonali e sfumature vocali, con cui la voce dell'attore simula i discorsi dei familiari.
Rispetto ai testi di altri monologanti, occorre evidenziare un dato importantissimo per una ricerca come questa, che si pone l'obiettivo

[14] La scelta dell'oralità vernacolare del siciliano appare per Enia decisamente più marcata rispetto, per esempio, alla lingua dei romanzi di Andrea Camilleri. Il noto autore siciliano mira a una chiarezza comprensibile anche per chi è privo di competenze linguistiche specialistiche: quando ricorre alle voci di più spiccato rilievo idiomatico, spesso per stilizzare il parlato dei personaggi protagonisti dei suoi racconti, Camilleri usa una glossa esplicativa interna al testo, che spiega la voce siciliana, con l'equivalente in italiano o con l'esplicitazione metalinguistica.

di evidenziare le interferenze tra la scrittura e il parlato: la scrittura assimila le modalità e le forme dell'oralità, al punto tale da eludere ripetutamente le regole tradizionali della morfosintassi, così come si sono codificate nella scrittura. Sulla pagina si instaura infatti una fortissima tensione fra *gramma* e *phoné*, dal momento che sul piano grafico l'imitazione e la riproduzione del flusso parlato presenta un fitto repertorio di fenomeni tipici dell'oralità, che, così come sono presentate nel testo, appaiono di fatto fuori dalla correttezza normativa. Si va dall'impiego dei vari caratteri tipografici (per esempio, il corsivo, il maiuscolo), per evidenziare, con un uso espressionistico e anomalo, le diverse *nuances* dei sopratoni della voce e delle sue sottolineature, fino all'estremistica rappresentazione del silenzio e dello stupore (come nell'esempio: *mi venne un'idea MI-CI-DI-A-LE…*, p. 38).

In particolare, l'impiego oralizzante delle forme grafiche e dei segni diacritici viene adottato per significare, sulla pagina scritta, la varia e vasta fenomenologia del rumore e dei toni della voce, non sempre risolvibili con scelte di carattere tipografico (Stefanelli, 2006: 154-155). Tale uso investe peraltro anche il sistema dei segni interpuntivi, essendo questo orientato a simulare sulla pagina i ritmi e i tempi dell'escursione melodica della voce attoriale. Più nello specifico, il modulo dei puntini sospensivi – icona grafica e chiave tonale della finzione della voce, usati per tradurre sulla pagina la frantumazione della voce enunciante (Mortara Garavelli 1985: 70) – serve a riportare e "significare" sulla pagina gli aspetti intonazionali della voce.

Per tutti questi fattori il monologo diventa un oggetto davvero interessante da esaminare per l'ostentato ricorso, in altissime occorrenze, a moduli compositivi e vari artifici grafici, usati per rendere visivamente le diverse possibilità tonali della voce, i chiaroscuri del processo enunciativo, o le irregolari variazioni della *phoné*, che entrano in azione durante l'esecuzione scenica.

A tal proposito, si riportano qui di seguito alcuni segmenti testuali, che esemplificano questo particolarissimo processo di interferenza tra la scrittura e l'oralità. Negli esempi che seguono:

Paulo Roberto Falcao, *accarezzò della carezza più gentile* (p. 43);
'a palla: *'un 'a pigghiò* (p. 25);
la sfera: *'a mancò* (p. 25), i segni diacritici della virgola e dei due punti sono usati per marcare evidentemente il *break* tonale, proprio del "modo" orale della lingua.

Inoltre intere battute risultano contrassegnate dall'allungamento della vocale, come queste :

GOOOOOOOOOOOOOOOOOOOOOOOOOOOOLLE! (p. 35);
'un c'era nìiente da fare (p.32);

Con questi espedienti grafici si tenta di riprodurre sulla pagina gli scarti tonali della voce, quasi a testimoniare che soltanto il ricorso alle risorse tipografiche consente di riparare a una *manque* costituzionale della scrittura, soprattutto lì dove non è possibile ricorrere a indicazioni sulle modulazioni della voce con le opportune didascalie, o sopperire altrimenti per mezzo del vasto repertorio dei modi di dire (i diversi *verba dicendi*), con cui spesso il narratore specifica le diverse modalità delle battute, simulando le intonazioni della voce. In questo senso, ancora più interessante diventa l'impiego di veri e propri procedimenti iconici, che riproducono le interruzioni, le pause e la frantumazione del parlato, ottenute attraverso un uso anomalo ed espressionistico dello spazio grafico, come testimonia l'esempio qui riportato:

quando

all'improvviso

dal nulla

da dietro la schiena d'u difensore brasiliano ca si chiama Oscar
appare una maglia azzurra con dentro un giocatore magro magro magro, nùmmero 20 nna schiena
ca si chiama: Paolorossi nato a Prato (p. 29).

Si riportano a questo punto i risultati della schedatura relativa agli aspetti morfosintattici del monologo; sul piano del lessico, si registrano alte occorrenze di sicilianismi: *si talìa intorno* (p. 28); *purtroppamente* (p. 29); *a me casa*; *me zio Peppe* (p. 29); *picciò* (p. 30); *talìa a Falcao* (p. 33); *lo quartìa* (p. 34); *abbannìamo come dei pazzi* (pp. 45); *macari* (p. 41); *minchia* (p. 46). Ancora, numerosi appaiono i metaplasmi popolareschi, *scancellato* (p. 32); o i malapropismi *Ucrania* (p. 36); *avversaio* (per *avversario*) (p.28); *armalo* (per *animale*) (p. 28).

In molte espressioni appare evidente l'uso del verbo in chiusura di periodo, come negli esempi che seguono: *troppo forti sono […] troppo scarsi siamo* (p. 32); *'u sport: vietatissimo è* (p. 36); *un pendolo pare* (p. 46). O anche l'opzione del passato remoto a scapito del passato prossimo, in particolare nel discorso diretto, come nell'esempio: *Ci fu pure una squadra… ca fu…*(p. 36).
Numerosi sono i raddoppiamenti fonosintattici, per imitare sulla pagina scritta un fenomeno proprio del parlato, diatopicamente marcato: *dribbla* (p. 28); *càrrica* (p. 28); *direzziòne* (p. 29); *imbattìbbili* (p. 30); *raggione* (p. 31); *impossibbìle* (p. 34); *cuggìno* (p. 37); *terribbìle* (p. 47); diversi casi di accusativo preposizionale: *Ammazza a cereo, spaccaci i gambe, pòtalo, pòtalo*" (p. 45); *ma vediamo a Bruno Conti* (p. 47).
Nel racconto appare evidente l'uso esteso del *ci* ridondante: *in mano ci ha una nazzionàle* (con raddoppiamento fonosintattico) *senza filtro* (p.31); *la palla scende rapace e / ci passa tra le braccia al portiere* (p. 35); *colpirà la palla di esterno sinistro dandoci al pallone* (p. 51).
Tratti caratteristici della ridondanza, propri dell'espressività del parlato, sono anche le forme della ripetizione, come, per esempio, la pausata *geminatio* degli aggettivi o l'iterazione a contatto con valore superlativo: *vesciche chìne chìne* (p. 36); *per portarci aggrìtta aggrìtta in paradiso* (p. 35); *ci passerebbe giusto giusto un pallone* (p. 45); *'u primo primo calcio d'angolo per l'Italia* (p. 47); *corre corre corre* (p. 45); *un giocatore magro magro magro* (p. 24).
Si riscontra un uso estensivo del deittico con ripetizione enfatica: *ma la palla chìdda del tiro… chìdda… chìdda* (p. 25). Si evidenziano inoltre casi di frase foderata, o epanalessi, nella tipica struttura a occhiale (detta anche "struttura a cornice"): *40 anni… avìemu in porta un portiere di 40 anni!* (p. 27); *bastava perdere, bastava* (p. 40).
Non mancano infine attestazioni di frasi con ordine marcato dei costituenti: *e la partita la addormentano. La anestetizzano* (p. 39); *ci sono giocatori che appena inizia 'a partita la sudano tutta la maglia* (p.42).

6 Conclusioni. Primi risultati della ricerca

La ricerca evidenzia che nel corpus analizzato si verifica un ispessimento della filigrana dialettale, nonostante alcuni testi siano più

marcati da opzioni dialettali (Celestini, Enia), mentre in altri domini, in prevalenza, un'oralità non idiomatica (Paolini).
In effetti, l'immediatezza comunicativa, l'intensità espressiva, la drammatizzazione della narrazione si ottengono con varie strategie e diversi mezzi linguistici: per Paolini si registra una resa più plastica e flessibile dell'italiano colloquiale, con poche inflessioni dialettali; mentre, per Celestini, il registro narrativo si orienta verso l'italiano popolare e la dialettalità del romanesco; per Enia, infine, le opzioni linguistiche si attestano decisamente sull'italiano regionale siciliano.
In conclusione, si presentano, qui di seguito, alcuni tratti comuni ai monologhi analizzati, con l'analisi di elementi relativi agli aspetti del lessico, della sintassi e della testualità:

innanzi tutto, occorre evidenziare che la variazione diatopica investe in particolare la fonetica e il lessico; le *performances* dei monologanti appaiono decisamente lontane dal modello, adottato dal teatro di prosa, dello standard parlato, o "parlato da laboratorio", esemplato sul cosiddetto «fiorentino colto emendato» (Galli de' Paratesi, 1984: 72);
per il livello del lessico, si registra l'occorrenza di numerosi colloquialismi e parole attinte dal vocabolario di base. Sullo sfondo di una dialettalità prevalente, non si verifica un'alternanza rigida tra italiano e dialetto, quanto, piuttosto, si evidenzia un'alta frequenza di intersezioni, come malapropismi, forme analogiche e diversi tipi di paronimia. Questo fenomeno (almeno per Celestini ed Enia) documenta un aumento della presenza dell'oralità dialettale sulla scena, che appare sicuramente più marcato rispetto, per esempio, al teatro di Eduardo De Filippo, lì dove la varietà propriamente dialettale viene recuperata come «codice della tradizione (in particolare di quella gastronomica: *annecchia* 'carne di vitellone', *peppiato* 'bollito a fuoco lento', *tiano* 'pentola di terracotta') o, più spesso, come registro tipico dell'impeto emozionale, forma espressiva adatta a soddisfare un bisogno di comunicazione sincera e intensa» (Trifone, 2000: 102-103)[15];
l'estrema linearità dello svolgimento e la ripetizione delle stesse parole conferiscono al *narratum* (dei monologhi analizzati) un

[15] Questo aspetto conferma quanto rivelato per l'uso del dialetto a teatro da Giovanardi (2002: 405-438), che ha operato sistematici spogli linguistici sulla lingua usata nella scuola partenopea.

andamento ripetitivo, soprattutto per l'iterazione degli stessi lessemi anche a breve distanza. Domina pertanto la ricorrenza come fattore della coesione testuale, poiché frequente appare il ricorso alla figura dell'*adiectio*, soprattutto all'anafora e all'anadiplosi. La coesione testuale è affidata in sostanza alla ripetizione, meccanismo frequente non solo nel parlato spontaneo, ma anche nello stile formulaico, tipico delle narrazioni orali. Nello specifico, le serie anaforiche appaiono peraltro collegate dalla tecnica del cosiddetto *embedding*, o incastro di frasi subordinate, introdotte spesso dalla congiunzione *che*, con valore prevalentemente relativo (soprattutto nel monologo di Celestini). In particolare tutto il monologo di Davide Enia è costruito con tecniche che ricordano lo stile formulaico, tipico delle narrazioni orali; con una notevole ridondanza si ripetono gli stessi lessemi, anche a breve distanza, come negli esempi che seguono: più volte ritorna un sintagma, come un vero e proprio *refrain*, *Paolorossi nato a Prato* (p. 25); o costruzioni anaforiche con sequenze così scandite:

dalla posizione di collo e spalle
dalla curvatura della schiena
dall'angolazione di bacino ginocchio caviglia
dalla parte del piede che impatta la sfera di cuoio
dalla potenza e dall'armonicità da immettere in
ogni singolo movimento (p.46);

sotto il profilo sintattico e interpuntorio, la tendenza più evidente è quella della misura breve del periodo. Ciò si traduce, sul piano testuale, nell'adozione di frasi brevi o brevissime e di un periodare franto e spezzato. In queste scelte influisce anche una funzione prettamente connotativa, dal momento che il monologo si distingue per un uso del ritmo nervoso e martellante delle frasi singole, frutto di scelte di tipo espressivo. Di fatto, anche la punteggiatura si caratterizza per l'affiorare di usi che riflettono il parlato: spesso è frequente l'uso antigrammaticale della virgola, posta tra elementi strettamente connessi dal punto di vista logico-sintattico, come il soggetto e il predicato verbale, a segnalare le pause tonali dell'oralità (per esempio, nel monologo di Enia: *perché sta preparando il tiro, Falcao*, p. 46).

In generale, sul piano sintattico prevalgono dunque schemi rapidi e lineari, con andamento paratattico e spesso coordinazione asindetica. Ovviamente, questo aspetto formale deriva anche da scelte contenutistiche, considerata la semplicità di svolgimento dei fatti

narrati e la visione *d'en bas* o punto di vista abbassato, adottato in molti monologhi;
nella fitta partitura oralizzante dei racconti si riscontra (soprattutto nel confronto tra testo a stampa e testo recitato) un aumento notevole dei moduli deittici, legati a tratti espressivi o anche mimico-gestuali. I fenomeni dell'indessicalità – per la forte implicazione mimico-gestuale, che li connota – indicano la ricerca dell'intensificazione espressiva, con l'uso delle forme marcate e insistite della deissi, la cui adozione appare funzionale alle esigenze di impatto comico-scenico;
è presente spesso una sottile trama polifonica, con la simulazione del discorso diretto, usato con evidente fine mimetico e con l'intento di stilizzare il parlato informale;
nell'esecuzione scenica, alle pause di esitazione si cerca di rimediare con riempitivi fàtici e segregati vocalici;
infine, l'intensità espressiva e la drammatizzazione dell'atto performativo si ottengono con varie strategie: per esempio la *geminatio* enfatica degli aggettivi; o le interiezioni, indicatori di empatia, che permettono di ancorare maggiormente il discorso al contesto pragmatico, mostrando così la forza performativa dell'atto enunciativo.

Si può dire, in conclusione, che la combinazione e la ricorsività dei tratti elencati costituiscono un dato caratteristico del parlato monologico del *teatro narrazione* e presentano elementi di compattezza e uniformità linguistica tali, da dimostrare sempre di più l'avvicinamento del testo teatrale ai *realia* dell'oralità.

Riferimenti bibliografici

Testi teatrali e narrativi

Celestini, A. (2006) *Scemo di guerra. Il diario*. Torino: Einaudi.

Enia, D. (2005) *Italia-Brasile 3 a 2*. Milano: Ubulibri.

Paolini, M. and Vacis G. (2008) *Vajont, 9 Ottobre '63. Orazione civile*. Torino: Einaudi.

Saggi linguistici e critici

Ariani M. and G. Taffon (2001) *Scritture per la scena. La letteratura drammatica del Novecento italiano*. Roma: Carocci.

D'Achille, P. (1990) *Sintassi del parlato e tradizione scritta della lingua italiana. Analisi dei testi e tradizione scritta dalle origini al secolo XVIII*. Roma: Bonacci.
– (2001) Parole in palcoscenico: appunti sulla lingua del teatro italiano dal dopoguerra a oggi, in M. Dardano, A. Pelo and A. Stefinlongo (eds) *Scritto e parlato. Metodi, testi e contesti*, Atti del Colloquio internazionale di studi (Roma, 5-6 febbraio 1999), Roma: Aracne, pp. 81-219.
Ferrone, S. (1992) La drammaturgia consuntiva, in J. Jacobelli (ed.) *Non cala il sipario. Lo stato del teatro*, Bari: Laterza, pp. 97-114.
Folena, G. (1991) *Il linguaggio del caos. Studi sul plurilinguismo rinascimentale*. Torino: Bollati Boringhieri.
Galli de' Paratesi, N. (1984) *Lingua toscana in bocca ambrosiana. Tendenze verso l'italiano standard: un'inchiesta sociolinguistica*. Bologna: Il Mulino.
Giovanardi, C. (2002) Plurilinguismo e antirealismo nel teatro napoletano dopo Eduardo, in M. Savini (ed.) *Studi di Italianistica per Maria Teresa Acquaro Graziosi*, Roma: Aracne, pp. 405-38.
– (2006) Il romanesco nel teatro contemporaneo, in N. Binazzi and S. Calamai (eds) *Lingua e dialetto nel teatro contemporaneo* (Atti della Giornata di Studio, Prato, Ridotto del Teatro Metastasio, 12 marzo 2004), Università degli studi di Firenze Quaderni del Dipartimento di Linguistica – studi 6, Firenze: Unipress, pp. 107-28.
Guccini, G. (2005) *La bottega dei narratori. Storie, laboratori e metodi di: Marco Baliani, Ascanio Celestini, Laura Curino, Marco Paolini, Gabriele Vacis*. Roma: Dino Audino.
Grignani, M. A. (2007) *Novecento plurale, Scrittori e lingua.* Napoli: Liguori.
Jakobson, R. (1966[1963]), *Saggi di linguistica generale*. Milano: Feltrinelli.
Lavinio, C. (1990) *Teoria e didattica dei testi*. Firenze: La Nuova Italia.
Lezza, A. (2005) *Vincenzo Pirrotta, N'gnanzou'. Storie di mare e di pescatori*. Salerno: Plectica.
Lucia, C. (2008) "Fabula, ficta, fabula agenda, fabula acta": il parlato nella transcodificazione dalla pagina alla scena, in M. Pettorino, A. Giannini, M. Vallone and R. Savy (eds) *La comunicazione parlata*,

Atti del Congresso Internazionale (Napoli 23-25 febbraio 2006), Napoli: Liguori, pp. 859-93.
Meldolesi, C. and G. Guccini (2004) 'Editoriale. L'arcipelago della nuova "performance epica"'. *Prove di drammaturgia* I, pp. 3-4.
Mortara Garavelli, B. (1985) *Ricognizioni. Retorica, grammatica, analisi dei testi.* Napoli: Morano.
Nencioni, G. (1976) *Parlato-parlato, parlato-scritto, parlato-recitato* [ora in *Id.* (1983) *Di scritto e di parlato. Discorsi linguistici*, Bologna: Zanichelli, pp. 126-79].
Petrilli, S. (1999) Traduzione e traducibilità, in S. Gensini (ed.) *Manuale della comunicazione. Modelli semiotici, linguaggi, pratiche testuali*, Roma: Carocci, pp. 419-49.
Platone (2000) *Fedro*. Milano: Bompiani [a cura di G. Reale].
Sabatini, F. (1985) L'«italiano dell'uso medio»: una realtà tra le varietà linguistiche italiane, in G. Holtus and E. Radtke (eds) *Gesprochenes Italienisch in Geschichte und Gegenwart*, Tübingen: Narr, pp.154-84.
Segre, C. (1984) *Teatro e romanzo. Due tipi di comunicazione letteraria.* Torino: Einaudi.
Stefanelli, S. (1987) Come parla il teatro contemporaneo, in Accademia della Crusca, Incontri del Centro di studi di grammatica italiana, *Gli italiani parlati. Sondaggi sopra la lingua di oggi*, Firenze: Accademia della Crusca, pp. 247-63.
– (2006) *Va in scena l'italiano*. Firenze: Cesati.
– (2007) I linguaggi del teatro di narrazione, in P. Puppa (ed.) *Lingua e lingue nel teatro italiano*, Roma: Bulzoni, pp. 331-54.
Trifone, P. (1994) L'italiano a teatro, in L. Serianni and P. Trifone (eds) *Storia della lingua italiana II. Scritto e parlato*, Torino: Einaudi, pp. 81-159.
– (1995) Una maschera di parole. La commedia fra grammatica e pragmatica, in M. Dardano and P. Trifone (eds) *La sintassi dell'italiano letterario*, Roma: Bulzoni, pp. 139-279.
– (2000) *L'italiano a teatro. Dalla commedia rinascimentale a Dario Fo*. Pisa-Roma: Istituti editoriali e poligrafici internazionali.

L'improvvisazione in ottava rima in Toscana

Una pratica di comunicazione "solenne"

Grazia Tiezzi
LIAS/IMM - EHESS
gratie61@hotmail.com

Abstract

The Ottava Rima *(the eight-line rhyming stanza abababcc) improvisation is a traditional poetry style that has survived from the Middle Ages to our own day in central Italy. This particular oral tradition has never interested linguistic and sociolinguistics experts. It has simply been considered as one example of the vast repertoire of Italian folk poetry. For this reason,* Ottava Rima *improvisation has been the object of analysis within literature and ethnomusicological studies. Such analyses have focused on the skills and the virtuosity of a few men, the 'extemporary' poets, capable of improvised performances by singing lines of verse to simulate verbal duels. However, my personal approach to the* Ottava Rima *improvisation and to its poetics, is oriented towards researching the type of social relationships that emerge between the participants through highly conventional interactions they build together during the dialogic performance. I observe the procedural aspects and I try to analyze the stylistic patterns displayed in the actual performance. I re-examine this type of poetic folk art using a holistic approach, but applied to the limited ethnographic context that corresponds to the area of the Tuscan Maremma. This approach implies establishing a possible convergence between the phenomenology of dialogic poetry improvisation and exchanges that occur in spoken language. I draw attention to the specific nature of this traditional poetic art by treating it as a practice of 'marked' communication. I propose to attribute* Ottava Rima *improvisation with the status of 'solemn' communication, which conserves the memory of a jurisprudential model of face-to-face verbal interaction involving two activities: the encounter and the dispute with the Other. In fact, in this verbal art, the intrinsic traits of multimodal expression (sung – versified – rhymed) seem to govern the procedures of speech transaction, and also seem to intervene as local resources in the practical management of conflictual interaction that follows precise regimes. According to the interpretation key that I propose for discussion, in the improvised, dialogic poetry, the individual virtuosity of the poets must negotiate within a complex framework, both expressive and praxeologic, which has a mediatory role, and whose built-in properties reveal an ethics of the dialogue. My work finally suggests a possible relationship between this oral tradition and the socio-economic activity of transhumance practiced by the communities of shepherds who moved from the Tuscan-Emilian Apennines to the Tuscan Maremma.*

0 Introduzione

Nel mio intervento vorrei esporre alcune brevi riflessioni che sono il risultato ancora provvisorio di una ricerca sull'improvvisazione poetica in ottava rima in Toscana. Tale ricerca è l'oggetto di una tesi di dottorato in scienze del linguaggio che sto svolgendo presso il LIAS/IMM (Linguistique Anthropologique Sociolinguistique/Institut Marcel Mauss) – Ecole des Hautes Etudes en Sciences Sociales.
L'improvvisazione in ottava rima è una tradizione orale molto antica ancora praticata con modalità simili nelle regioni dell'Italia centrale quali la Toscana, il Lazio e l'Abruzzo ma è presente e viva anche in Sardegna. La mia tesi tenta di rivalutare la pratica di questa forma poetica estemporanea attraverso una nuova prospettiva d'analisi che fa riferimento al quadro teorico della Linguistica interazionale e della Pragmatica socio-cognitiva. Questa tradizione culturale è osservata all'interno di un contesto etnografico preciso e la descrizione dei suoi aspetti praxeologici è condotta adottando un approccio etnometodologico.

1 Corpus

Le basi empiriche a cui si àncora la proposta di una nuova prospettiva d'analisi sono costituite da una lunga raccolta di dati che mi ha permesso di costituire un archivio personale di registrazioni audio ed audiovisive[1]. Dal 2004 ho cominciato a videoregistrare molte gare d'improvvisazione in Toscana e nel Lazio ed ho deciso di focalizzare la mia analisi sulle improvvisazioni poetiche toscane della Maremma grossetana e della provincia di Arezzo. Il corpus di base su cui ho lavorato raccoglie le gare d'improvvisazione realizzate in Maremma da poeti toscani provenienti dalle province di Grosseto, Arezzo, Pisa, Firenze, Livorno e Pistoia. Rispetto alla diffusione di questa pratica poetica, la scelta di concentrare l'attenzione su un repertorio che esclude di fatto le competizioni laziali ed abruzzesi, può sembrare molto riduttiva ed è infatti volutamente parziale rispetto all'estensione geografica del fenomeno. Non è parziale invece se si considera questa

[1] I materiali a cui si fa riferimento nel presente contributo possono essere richiesti a Grazia Tiezzi al seguente indirizzo: Loc. Casone, 50 – Alberoro 52048 Arezzo (Italia), e-mail: gratie61@hotmail.com.

scelta come un parametro per rappresentare la fenomenologia contemporanea di ciò che resta di questa forma d'improvvisazione in un'area socio culturale identificabile grosso modo con la Maremma grossetana di questi ultimi venti anni. I punti di riferimento principali per la costituzione del corpus sono i dati relativi all'appuntamento annuale di "Poesia Estemporanea" di Ribolla (Grosseto) e le riunioni di poesia del Circolo di Pianizzoli (Grosseto). Questa scelta è motivata dal fatto che le *performance* poetiche toscane sono accomunate da alcuni elementi che riguardano le condizioni di produzione e di svolgimento delle competizioni poetiche. Non ho ritrovato le stesse caratteristiche nelle competizioni del Lazio (Tolfa, Bacugno, Amatrice) alle quali ho avuto la possibilità di assistere. Riassumo di seguito gli aspetti che ho considerato peculiari delle competizioni toscane:

l'assenza di una giuria esterna giudicante la gara poetica;
l'assenza di indicazioni restrittive imposti dall'esterno riguardo la predeterminazione di elementi formali aggiuntivi – per esempio il primo verso di strofa o il tipo di rima – o riguardo la durata della *performance*;
la costante applicazione dell'obbligo di ripresa di rima;
l'assenza di un vincitore della gara identificato da una giuria esterna;
la presenza attiva del pubblico nell'indicazione dei temi delle sfide poetiche che di fatto determina anche il numero dei poeti che si devono affrontare nelle *performance*.

Di fatto lo svolgersi delle gare poetiche toscane non sembra essere vincolato da obblighi imposti dall'esterno da parte di una giuria. Questo elemento porta alla mancanza di un vincitore della gara identificato da esperti esterni ai partecipanti della performance. Ho ritenuto che questo aspetto fosse rilevante per accedere alle regole endogene della pratica dell'improvvisazione in ottava rima. La loro identificazione potrebbe far emergere le tracce di un possibile funzionamento auto-regolativo soprattutto riguardo l'esecuzione delle *performance* di tipo oppositivo. Inoltre ad una prima osservazione mi è sembrato di cogliere una certa continuità strutturale tra lo svolgimento delle competizioni della manifestazione di Ribolla e le occasioni d'improvvisazione nelle riunioni poetiche più informali che ho potuto osservare direttamente.

2 Lavori precedenti

Il mio interesse per questa tradizione orale si concentra sulle improvvisazioni dialogiche in *ottava rima* ed esclude tutto il repertorio dei canti estemporanei di tipo narrativo/ epico monologico, a lungo privilegiati dagli studi precedenti. Di fatto l'improvvisazione poetica dialogica non ha mai suscitato l'interesse dei linguisti, esclusi alcuni studi in ambito dialettologico, ed è, ancora oggi, un oggetto di studio trascurato in particolare dai sociolinguisti. L'analisi più approfondita in questo senso resta quella prodotta in ambito musicologico da Maurizio Agamennone, il primo ad essersi veramente occupato di mettere in relazione gli aspetti della forma dell'espressione orale, in particolare per lui quelli melodici, con le interazioni costruite dai poeti sfidanti durante le gare d'improvvisazione. Nei suoi saggi *Maestri del tempo, maestri della voce* (Agamennone, 1999) e *Modi del contrasto in ottava rima* (Agamennone, 2002), l'etnomusicologo ha approfondito lo studio del contrasto poetico ed ha messo in evidenza alcune caratteristiche dell'interazione che ha definito come "cooperazione antagonistica"[2]. Per l'autore, che si rifà a Zumthor (1984), la "stilizzazione della disputa" ovvero la sua "rigida regolamentazione" definisce il quadro della cooperazione e il virtuosismo individuale dei poeti permette l'espressione dell'antagonismo. La struttura formale è orientata verso procedure cooperative mentre l'abilità soggettiva nel manipolare la struttura è orientata verso l'espressione dell'opposizione all'avversario.

3 Prospettiva di analisi

Il mio lavoro si inscrive nella linea aperta da Agamennone, però, nel mio studio, i fenomeni rilevati della sua analisi musicologica sono in parte rimessi in discussione e vengono trattati anche con gli strumenti descrittivi della prosodia del parlato. Inoltre tento di praticare una analisi integrata della pragmatica dell'interazione degli scambi dialogici con la forma *multimodale* dell'espressione delle ottave

[2] L'autore segnala di aver ripreso tale denominazione dalla definizione data da Alessandro Portelli nella conferenza *Il ruolo dell'oralità nella poesia*, tenutasi ad Artena nel 1996. Cfr. nota 20 in Agamennone (1999: 147).

improvvisate, costituita dal canto versificato e rimante. Da alcuni risultati ottenuti praticando una descrizione *intermodale* delle convenzioni della melodia, della metrica e delle rime, emergerebbe un'articolazione più complessa tra i livelli formali della strofa[3]. La complessità è dovuta al parziale non isomorfismo rilevato tra le differenti unità della forma dell'espressione orale. La parziale non convergenza sistematica tra le unità intonative dei moduli melodici del canto e le unità metriche della versificazione rimante rende la strofa cantata una struttura a morfologia pluri-modulare disponibile a segmentazioni variabili e non sembra essere soltanto un unico e stabile dispositivo strutturale che permette la collaborazione tra gli improvvisatori. Con il mio approccio, inoltre, intendo riconsiderare la *natura* di questa pratica d'improvvisazione in relazione alla comunicazione parlata ordinaria. Applicando le nozioni e gli strumenti della Linguistica interazionale, dell'Analisi della conversazione e della Pragmatica linguistica, tento di identificare i tratti costitutivi delle interazioni fortemente convenzionalizzate emergenti dagli scambi poetici.
Lo scopo è, da un lato, quello di stabilire un possibile accostamento tra le improvvisazioni dialogiche in ottava rima e gli scambi comunicativi della conversazione ordinaria e, dall'altro, quello di stabilire la specificità della differenza tra i vincoli dell'esecuzione degli scambi poetici e le forme di scambi della comunicazione quotidiana. Se da un punto di vista empirico la loro differenza sembra evidente, la specificità della loro distinzione a livello socio pragmatico è ancora tutta da capire. In altri termini cerco di *interpretare* in senso sociolinguistico lo scarto che c'è tra questi due modi di esistenza dello scambio dialogico.

4 Ipotesi interpretativa

Seguendo questa metodologia l'ipotesi che mi sembra possibile avanzare per il momento, è quella di considerare l'improvvisazione

[3] Questo tipo di descrizione segue un principio di *sincretismo della pertinenza distintiva* che si basa sul rilevamento di una proprietà della morfologia della strofa cantata quale la parziale asimmetria progressiva tra le unità dei diversi livelli formali. Ho approfondito questo aspetto nella mia tesi di dottorato che è in fase di ultimazione.

dialogica in ottava rima, non tanto una forma di poesia popolare quanto una pratica discorsiva "marcata" rispetto a quella della lingua parlata ordinaria. L'improvvisazione poetica dialogica costituirebbe un sistema di comunicazione secondario e specializzato a cui si può attribuire lo *statuto performativo* di comunicazione "solenne". Con questa attribuzione, ancora approssimativa, intendo identificare uno stile o una modalità di organizzazione del discorso, rigidamente regolamentata e virtuosa, investita in modo consuetudinario, di una *funzione istituzionale* e *sociale*. Le proprietà multimodali dell'espressione e l'organizzazione degli scambi comunicativi improvvisati in ottava rima avrebbero una loro agentività. Agirebbero come *strumento di governo* di alcune interazioni verbali "a rischio" e si farebbero garanti del successo di specifiche relazioni intersoggettive "sensibili" per la stabilità dell'ordine pubblico. Con interazioni "a rischio" mi riferisco più precisamente a due macro aree dell'interazione verbale quotidiana che sono genericamente indirizzate, in un caso a stabilire il *contatto* con l'Altro, nell'altro a realizzare il *confronto* con l'Altro. Le proprietà intrinseche della pratica di quest'arte verbale, farebbero emergere la sua specializzazione che consiste nel rappresentare e nell'esibire pubblicamente delle precise modalità dialogico-pragmatiche finalizzate a *stabilire con successo* una "prossimità" ed a *gestire con successo* una "prossimità conflittuale". Di fatto la fenomenologia dell'improvvisazione dialogica toscana declina solo dei quadri di attività riferibili all'incontro e all'antagonismo. Entrambe queste aree d'attività possono facilmente risolversi con l'insuccesso dello scambio comunicativo e sono quindi a rischio di fallimento della relazione sociale. Il *contatto* può essere respinto o differito e il *confronto oppositivo* può trasformarsi in conflitto.

5 Proprietà di coordinazione

Ciò detto, mi preme mettere in evidenza che lo statuto di *solennità* attribuibile alla pratica d'improvvisazione non deriva da una investitura o da ruoli istituzionali extralinguistici degli interlocutori. Ricordo che questo è uno dei presupposti della *condizione di felicità* dei performativi espliciti istituzionali nella definizione di Austin (1962). I poeti improvvisatori erano e sono spesso dei contadini, degli artigiani, dei pastori, dei boscaioli, degli operai, persone comunque

senza nessuna carica istituzionale all'interno della comunità. Secondo la mia analisi, l'*autoritas* sociale di questa pratica poetica sembra derivare piuttosto dalle proprietà che assume la coordinazione dell'attività durante l'improvvisazione. Ha quindi un carattere riflessivo ed endogeno, esibito dalla *prassi* condivisa e incarnato nella *tecnica* esperta del canto versificato degli improvvisatori. Infatti l'esecuzione delle *performance* poetiche improvvisate avviene secondo dei canoni consuetudinari che prevedono dei vincoli procedurali ed espressivi molto rigidi. Il rispetto di tali vincoli costituisce anche la base della intercomprensione e dell'accettazione dell'attività in corso da parte dei co-partecipati. Ne costituisce quindi la *condizione di felicità,* cioè opera in modo che tali tipi di scambio si compiano con successo. In questo caso faccio riferimento alla reinterpretazione in senso interazionale e conversazionale della *condizione di felicità* nell'interazione verbale e co-verbale faccia a faccia, esposto da Goffman (1986[4]):

> Tout cela m'amène à hasarder une définition de la condition de félicité qui se cache derrière toutes autres, d'une définition de la Condition de Félicité, à savoir toute disposition qui nous incite à juger les actes verbaux d'un individu comme n'étant pas une manifestation de bizarrerie.
> (…) c'est qu'il convient désormais de considérer les analyses syntactiques et pragmatiques comme décrivant empiriquement et en détail la façon dont nous sommes obligés de manifester notre santé mentale pendant les interactions verbales, que ce soit par la gestion des nos propres paroles ou par les preuves que nous donnons de notre compréhension de celles d'autrui (…) cela exige que nos dires s'appuient sur une certaine série de présuppositions et en évitent une autre. Répliquant à autrui, nous devons répondre non seulement à ses mots, mais aussi à sa pensée – afin qu'il puisse tirer parti à la fois de la scène locale et du monde plus vaste et plus distant de son expérience. Sur ce point, les sociologues n'ont pas été dernièrement d'un grand secours; ils n' ont fait que repérer la proposition et en donner des illustrations; mais pas la moindre analyse du tenu-pour-acquis; aucun cadre; pas la plus petite classification. En revanche, la répercussion des présuppositions sur la forme de surface et la prosodie des

[4] Una riflessione sulla definizione di "condizione di felicità" in Goffman è proposta in Fornel (1989).

> énonciations a été grandement éclairée par les linguistes, qui ont également étudié les pratiques d'interprétations de textes et des répertoires des réinterprétations. Mais l'importance du fait de se connaitre ou d'être intimes, de l'engendrement de la construction délibérée d'une biographie commune, du fait d'être ou non dans un état de parole, des divers localisateurs employés pour constituer un cadre aux énoncés qu'on veut être à mesure d'émettre succinctement, toutes ces questions critiques n'ont guère été étudiées. Derrière tous cela, liant ensemble tous ces thèmes, il y a la place socialement prescrite de ce qu'ont tient pour opération d'esprit. C'est une affaire de qui peut dire quoi à qui, dans quelles circonstances, avec quel préambule, sous quelle forme de surface, le tout sans être jugé insensé étant donné les relectures possibles. L'affaire de ce qui peut se dire tout en satisfaisant la Condition de Félicité (Goffman, 1986: 95-6).

Ora, per comprendere come la condizione di felicità sia soddisfatta nelle improvvisazioni poetiche è necessario fare riferimento al sistema di pertinenza degli scambi dialogici in ottava rima e avere accesso ad una serie di presupposizioni che compongono l'insieme dei tratti specifici di questo sistema comunicativo tradizionale e dei tipi di interazioni convenzionali che tale sistema permette di compiere. Tali presupposizioni costituiscono un "Common ground" condiviso tra i membri della "comunità di parola" (Silverstain, 1998) ovvero tutti i partecipanti ad una gara poetica compreso il pubblico. In particolare, le improvvisazioni poetiche tradizionali presuppongono una comprensione pratica tra gli interagenti che essi manifestano nell'esecuzione attraverso un tipo di coordinazione specifica. L'intercomprensione sull'attività segue una prassi dell'ordine del saper-fare espressivo che possiede delle precise proprietà deontiche con cui viene imposto un dover-fare dialogico. Di fatto le proprietà deontiche dell'espressione sembrano svolgere una forma di giurisprudenza e di governo dell'interazione orientando gli scambi verso un modello deontologico del dialogo ed una precisa concezione dell'interlocutore. In altre parole, il tipo specifico di coordinazione dell'attività comunicativa è la base della condizione di felicità per la corretta esecuzione tecnica dello scambio improvvisato, questa si àncora sulle proprietà deontiche della poetica e ne costituisce al tempo stesso la manifestazione. Così attraverso le sue proprietà intrinseche,

la forma poetica svolge la sua funzione sociale e investe questo "modo di esistenza" del dialogo di una solennità popolare.

6 Classi di *performance* dialogiche nelle gare d'improvvisazione toscane

L'organizzazione globale delle competizioni toscane segue una sceneggiatura tripartita costante. Una fase di apertura, una fase centrale ed una fase di chiusura. I poeti improvvisatori in ciascuna di queste tre fasi realizzano canonicamente delle *performance* improvvisate che corrispondono a delle denominazioni "native" precise la cui sistematicità è stata fino ad oggi ignorata dagli analisti. Seguendo invece la categorizzazione degli agenti, mi sono accorta che tali denominazioni corrispondevano ad una classificazione tradizionale relativa esclusivamente alle improvvisazioni dialogiche. Nella fase di apertura vengono realizzate le ottave dette "i saluti", nella fase centrale quelle dette "i contrasti" e nella fase finale le ottave dette "i congedi" o "i ringraziamenti". Nella Maremma grossetana inoltre è ancora oggi molto viva e diffusa la tradizione del Canto del Maggio, un rituale di questua, in cui vengono improvvisate delle ottave dette "il permesso". È facile notare come ogni tipo di *performance* identifichi anche un preciso *atto* di parola.
Dal punto di vista della costruzione interazionale, nelle competizioni emergono solo due classi di contesti d'interlocuzione. La prima riunisce le ottave di saluto e di congedo o di ringraziamento e di omaggio, improvvisate in apertura e in chiusura di una gara poetica, mentre alla seconda classe appartengono le ottave a contrasto improvvisate nella fase centrale: i duelli poetici tra due o più poeti avversari. Le due classi di contesto si diversificano rispetto ai quadri di partecipazione che sono mobilitati e ai formati di produzione pertinenti esibiti nella realizzazione dello scambio comunicativo in ciascuna delle due classi. Il contesto d'interlocuzione è un primo criterio distintivo per identificare le proprietà di coordinazione delle *performance* di saluto e di congedo come una classe di attività e le proprietà di coordinazione delle *performance* a contrasto con un'altra classe di attività. L'analisi integra questo primo criterio con i valori azionali compiuti nelle *performance* di ciascuna classe.

6.1 I Saluti

Propongo come esempio della prima classe di contesto d'interlocuzione la trascrizione di una breve sequenza di ottave di saluto (esempio n. 1), realizzata tra i poeti improvvisatori e il pubblico in apertura della manifestazione di "Poesia Estemporanea" di Ribolla del 1995.
Esempio n. 1: ottave di saluto, sequenza strofe 1-4, Ribolla, 1995[5].

P1: **Libero Vietti, strofa 1**
Due anni fa ricordo son venuto
Il tempo passa e ce lo ruba il buono
Anche stasera rinnovo il saluto
Mi dovete accettare come sono
Che quando un uomo si sente perduto
Che si decide ad abbandonare il trono
Io qui mi sento un fuoco quasi spento
Auguri a tutti e buon divertimento
PU : (applauso)
P2: Florio Londi, strofa 2
In questo giorno volontà mi sento
Di tornar con la vecchia poesia
Dove ci parla il cuore e il sentimento
E tutti quanti mette in armonia
Però oggi mi sento un lume spento
Da troppo tempo la passione mia
Ha cessato questa attività/che l'ha cessato
Chissà se i due terzi ho svalutato
PU : (applauso)
P3: Altamante Logli, strofa 3
Io come terzo mi son presentato
Vengo dal solo vicino a Florenzia
Dove l'ingegno umano è sviluppato
Dove il Rinascimento fe' partenza
E io qui stasera venni invitato
Per soddisfare le brame dell'udienza

[5] La trascrizione di esempio n. 1 si riferisce ad una sequenza di quattro ottave, eseguite durante la manifestazione annuale di "Poesia Estemporanea" di Ribolla (Grosseto) del 1995. La sequenza delle strofe di saluto comprende complessivamente diciannove ottave. I contributi successivi dei partecipanti sono indicati con le sigle in grassetto: P1, P2, P3, P4 per i poeti improvvisatori e con la sigla PU per quelli del pubblico. La registrazione è stata realizzata da Corrado Barontini e fa parte del corpus raccolto da Grazia Tiezzi.

Un saluto vi porto col mio cuore
Ai gentili signori e alle signore
PU : (applauso)
P4: Edilio Benelli, strofa 4
Chi si presenta fa l'agricoltore
Improvvisando davvero in questo istante
Un augurio voglio fa' di cuore
Ora mi provo se sarò bastante
Auguri ai signori e le signore
Io nel canto ne sarò badante
E il mio canto speriamo che frutti
Con la speranza d'accontenta' tutti
PU: (applauso)
(…)

Nella prima classe di *performance* esemplificata qui dalle ottave di saluto, il quadro di partecipazione è costante, la costruzione dialogica è diadica esclusiva e si realizza tra un locutore individuale, il poeta improvvisatore, ed un interlocutore collettivo, il pubblico. I poeti presenti hanno tutti lo stesso statuto di partecipazione, quello di locutori potenziali ed auditori che si succedono uno dopo l'altro, ma non sono mai interlocutori tra di loro. In questo quadro dialogico diadico, i partecipanti co-costruiscono uno scambio confermativo il cui valore pragmatico-interazionale è dato dalla conferma del consenso proiettato ed atteso dal poeta improvvisatore con la sua ottava, sull'interlocutore collettivo, cioè l'auditorio, rispetto ad un atto performativo esplicito di cui il pubblico è destinatario diretto esclusivo. I performativi rilevati si riferiscono tutti alla sfera dell'attività di contatto: un saluto, un omaggio, una richiesta di permesso, un congedo, un ringraziamento. Sia il contributo del locutore (l'ottava di saluto) sia quello dell'interlocutore (l'applauso) sono dal punto di vista semantico-pragmatico mono-azionali e mono-orientati. Il poeta tende verso la proiezione di un atto performativo di contatto e l'interlocutore verso la conferma della comprensione dell'atto e del consenso con l'attesa del locutore. Il formato di produzione pertinente è dato a livello di organizzazione locale, da una coppia adiacente a due turni, stabile, ed a livello sopra locale, dalla reiterazione della stessa coppia che conserva le stesse caratteristiche strutturali. Questa reiterazione organizza la sequenza degli scambi confermativi delle *performance* di apertura e di chiusura della competizione poetica. Il primo membro della coppia è la presa di

turno del poeta che realizza la sua ottava come T1, in cui c'è l'enunciazione di un atto performativo esplicito che si posiziona tendenzialmente nei versi del distico finale di strofa ovvero subito prima del PRT (punto di pertinenza transazionale) della coppia adiacente che è previsto e prevedibile. Il secondo membro della coppia è dato dall'applauso del pubblico come T2 che completa in modo preferito la coppia adiacente orientata al consenso verso il contatto proiettato con l'atto performativo del contributo del poeta. La coordinazione dell'attività conversazionale esibisce, attraverso l'incatenamento confermativo, il successo della conferma del contatto proiettato ed atteso dal destinatario. La modalità di coordinazione dello scambio in questo quadro d'attività rimanda ad una situazione d'enunciazione chiusa, reiterata rigidamente per ogni locutore potenziale. Un poeta non può prendere il turno di un altro poeta né può rivolgersi, in questa fase, agli altri poeti.

6.2 I Contrasti

L'altra classe di contesto d'interlocuzione riguarda le *performance* a contrasto. Propongo a titolo di esempio la trascrizione di una breve sequenza di ottave a contrasto tra due poeti avversari, Libero Vietti e Altamante Logli (esempio n. 2). La sequenza fa parte del contrasto poetico sul tema Tangentopoli e Magistratura che è stato richiesto dal pubblico durante la competizione di Ribolla del 1993.
Esempio n. 2: contrasto "Tangentopoli/Magistratura", sequenza strofe 1-6, Ribolla, 1993[6].

P1: Libero Vietti,Tangentopoli, strofa 1
Siccome io sono sempre un fortunato
Ci fecero il biglietto i dirigenti
Indovinate cosa mi è toccato
M'è toccato il biglietto alle tangenti

[6] La trascrizione di esempio n. 2 si riferisce ad una sequenza di strofe del contrasto Tangentopoli e Magistratura eseguito durante la manifestazione annuale di "Poesia Estemporanea" di Ribolla (Grosseto) del 1993. Il contrasto poetico comprende complessivamente dodici ottave. Nella trascrizione i contributi dei partecipanti sono indicati con le sigle in grassetto: P1 per il poeta Libero Vietti (Tangentopoli), P2 per il poeta Altamante Logli (Magistratura) e con la sigla PU per il pubblico. La registrazione è stata realizzata da anonimi e fa parte del corpus raccolto da Grazia Tiezzi.

È una cosa che ho sempre lottato
Contro questi brutali delinquenti
Però in questo momento è che mi arrendo
Co i' sangue e con il cuore li difendo
Pu : (applauso)
P2: Altamante Logli, Magistratura strofa 2
La rete di visura ora la stendo
Io tornerò colà sopra il potere
Nasce la Tangentopoli e comprendo
C'è una classe che 'un fece il su' dovere
Ora anche i pesci grossi io te li prendo
Come la va a finì lo vo' sapere
Io rammento lo sai son qui Di Pietro
E vado avanti e n' 'un ritorno indietro
PU: (applauso)
P1: Tangentopoli, strofa 3
Tu l'hai avuto te preciso il metro
Sento scorrere il sangue nelle vene
Prima che al mondo arrivasse Di Pietro
Noi industriali si andava tanto bene
Poi la magistratura ed il suo spetro
Te l'ha sparse l'Italia le catene
E ora lo devo dire a tutti quanti
Da quando c'è lui 'un si va più avanti
PU: (applauso)
P2: Magistratura, strofa 4
Si sbagliarono i nostri governanti
Anche se attacco un governo che c'era
Ti speculò troppo sugli ignoranti
E pe'ingannare i' popolo ebbe maniera
Ma la magistratura sta davanti
Bisognerà 'ngrandire ogni galera
Leva' questi imbroglioni dalla scena
Ogni carcer d' Italia resti piena
PU: (applauso)
P1:Tangentopoli, strofa 5
Di Pietro la tua mente è ben serena
E io invece appartengo a quei partiti
Dove te l'hanno fatta l'altalena
Che l'uno l'altro non son dei puliti
Ma te lo devo dire a mente piena
Accetto tutti quanti i vostri inviti
Ma anche fra i magistrati in dove lotti
Anche lì ce l'avete dei corrotti
PU: (applauso)
P2 :Magistratura, strofa 6

C'è gente da ufficio e da salotti
Ma guarda vedi qualche magistrato
Per l'onestà sincera d'omini adotti
Lo vedi con la vita gli'ha pagato
Per via dei ladri e brutti galeotti
Quanto sangue dell'arma si è versato
A questo non ci si dovèa arrivare
Si dovranno pentire e vergognare
PU: (applauso)
(…)

Nei contrasti si co-costruisce un'attività dialogica multi-partecipante e non solo diadica. Rispetto agli statuti di partecipazione assunti dai presenti, i poeti sfidanti sono sia interlocutori tra loro sia interlocutori del pubblico il quale assume e conserva lo statuto di interlocutore collettivo per tutta la durata della *performance*. Per questo è uno degli interlocutori possibili a cui il poeta può indirizzare la strofa o parte di essa. La modifica dell'orientamento partecipativo all'interno della strofa è usato dai poeti sfidanti a fini tattici per costruire l'efficacia pragmatica delle loro mosse, cioè per disegnare di volta in volta il sistema di alleanze e di opposizioni più efficace per lo sviluppo delle loro argomentazioni. In questo senso gli scambi a contrasto hanno un duplice orientamento sia confermativo sia disconfermativo. La costruzione dell'attività tende ad articolare, da una parte, la proiezione e l'attesa del consenso del pubblico, dall'altra, l'attesa e la proiezione del disaccordo verso l'avversario. In questo quadro multi partecipante la situazione d'interlocuzione è più aperta e la pianificazione pragmatica dell'attività diventa fortemente interattiva. Le strofe improvvisate reciprocamente dai due poeti sfidanti diventano delle prese di turno multi-turno e multi-azionali, in cui la struttura dei formati di produzione ha un carattere relativamente variabile.

6.2.1 Pianificazione pragmatica nei contrasti

All'interno dello spazio di parola di ciascun sfidante si possono individuare dei protocolli abbastanza regolari di pianificazione delle azioni. Di solito, fatta eccezione della prima ottava di un contrasto, nelle strofe si rilevano due piani d'azione costanti che comprendono due fasi successive ed ordinate di pianificazione dell'attività. Prima,

una fase di difesa subito seguita da una fase di offesa. Ognuna di queste fasi porta al compimento di alcune azioni protocollari anche queste tendenzialmente successive ed ordinate. Entrambe le fasi possono essere il luogo d'introduzione di movimenti tematici nella progressione argomentativa del confronto oppositivo. Dalla distribuzione tendenziale delle azioni emergono due modelli virtualmente attivabili di organizzazione pragmatica della strofa. Un modello prevede due azioni protocollari, la risposta nella fase della difesa e la pianificazione dell'attacco nella fase dell'offesa – compiuto, ad esempio, attraverso il lancio di un'accusa o di una minaccia in finale di strofa. Un altro modello invece prevede tre azioni protocollari, due nella fase di difesa cioè l'azione di contro attacco per replica immediata seguita dalla risposta, e un'azione nella fase dell'offesa, di nuovo l'azione di attacco. Questi valori azionali si posizionano in modo costante in alcuni versi cantati della strofa. Il contrattacco per replica immediata è realizzato regolarmente al verso 1, la risposta tendenzialmente ai versi 3-4 e l'attacco nel distico finale ai versi 7-8. In questo senso ogni ottava può essere considerata una presa di turno multi-turno, composta da formati sequenziali protocollari i cui valori azionali corrispondono alle posizioni specifiche di alcuni versi dell'unità strofica. I formati sequenziali con valore azionale si ancorano a degli indici di contestualizzazione di tipo prosodico-intontivo rilevabili nella struttura melodica della strofa cantata.

In quest'arte verbale anche le convenzioni della versificazione seguono un ordinamento sequenziale (Fornel and Tiezzi, 2008) nell'esecuzione interna della strofa, le più evidenti sono quelle che riguardano i vincoli fonici del sistema di rime dell'ottava toscana che prevede tre variabili foniche di rima, le fonie in A, in B e in C. La loro distribuzione è costante e vincolata dall'alternanza di A e di B, nelle parole-rima finali di verso della sestina iniziale e dalla ripetizione di C, nella rima baciata tra le parole-rima finali di verso del distico finale di strofa: AB AB AB CC. Le strofe dei contrasti toscani, inoltre, vengono eseguite con l'obbligo di ripresa di rima, dette anche "ottave incatenante". In questo modo i vincoli delle fonie del sistema di rime producono dei legami fonici tra due ottave successive realizzate dai poeti avversari. L'obbligo di rima impone la stessa fonia solo nei versi dispari della sestina della strofa successiva, perciò nelle ottave incatenate i primi tre versi dispari sono vincolati da legami fonici

interstrofici, i primi tre versi pari sono liberi dai vincoli fonici interstrofici e i due versi finali sono liberi da legami fonici per il poeta attuale ma vincolanti per il poeta successivo: AB AB AB CC/CD CD CD EE .
Se associamo i vincoli fonici tra le parole-rima con le azioni conversazionali, posizionate tendenzialmente in alcuni precisi versi rimanti delle ottave incatenate, scopriamo che ogni variabile fonica di rima indicizza un tipo diverso di azione. Possiamo allora trasferire le proprietà vincolanti delle fonie di rima, che riguardano la forma poetica, alle azioni conversazionali che riguardano la pianificazione pragmatica delle ottave a contrasto.

6.2.2 Forma poetica e forma dell'interazione

L'associazione tra prese di turno di parola, protocolli di pianificazione pragmatica e fonie della versificazione rimante risulta rilevante per caratterizzare le proprietà di coordinazione dell'attività oppositiva nell'interazione di un contrasto toscano. Da una parte, i vincoli fonici interstrofici impongono agli interlocutori delle restrizioni semantiche nelle scelte lessicali delle parole-rima (Tiezzi, 2008), dall'altra, sottopongono la pragmatica dello scambio dialogico a dei regimi di intenzionalità controllata. Infatti, in apertura di strofa, il poeta che prende la parola, è vincolato a rispettare un accordo fonico con il suo interlocutore avversario per realizzare l'azione di contro-attacco al verso 1, in variabile di rima ripresa. Poi dispone di una variabile fonica libera – che deve comunque essere nuova – per compiere l'azione della risposta al verso 4. Infine, per realizzare l'attacco finale al verso 8, deve utilizzare una fonia di rima nuova, non vincolata ma vincolante per il suo avversario che sarà obbligato a sua volta in apertura del suo turno di parola, a rispettare l'accordo fonico imposto dal poeta precedente. Ovvero, per poter esprimere la sua opposizione dovrà farsi carico della parola altrui accordandosi con la fonia di rima lasciata dal suo avversario. È importante notare che in questo dispositivo di vincoli fonici e dialogici solo l'azione iniziale di contro-attacco è vincolata dall'interlocutore e solo l'azione finale di attacco è vincolante per l'interlocutore. L'associazione tra poetica, struttura conversazionale e pianificazione semantico-pragmatica mette in evidenza i tratti di una deontologia che i poeti sono obbligati a

rispettare per coordinarsi nell'attività e per adempiere alla corretta esecuzione degli scambi poetici a contrasto.

7 Regimi di restrizioni

Nella pratica attualizzata dell'improvvisazione dialogica si ritrovano dei protocolli procedurali condivisi dai partecipanti (poeti e pubblico) che sono esibiti attraverso la realizzazione metodica di routine di azioni. Questi protocolli stabiliscono un doppio regime di regole, quelle per *l'accesso alla parola* e quelle per la produzione delle *azioni comunicative possibili* con l'accesso alla parola. Tali protocolli sono il risultato di precise restrizioni nell'organizzazione conversazionale rispetto al dialogo ordinario. Per *l'accesso alla parola* le restrizioni riguardano il *sistema della presa di parola* – quando si deve parlare e quanto si deve parlare. Infatti, l'ordine di allocazione dei turni, stabilito all'inizio normalmente per autoselezione, resta stabile fino alla fine della *performance*, inoltre non c'è la possibilità per gli improvvisatori di saltare il proprio turno e di rimanere in silenzio. Anche la durata del turno è prevedibile poiché è data dalla strofa d'ottava – prototipo della stanza nella metrica italiana (Beltrami, 1994) – che è la misura metrica usata in tutti gli scambi improvvisati toscani. È importante notare che nella conversazione quotidiana l'interlocutore ha la possibilità di rimanere in silenzio e non c'è restrizione sulla durata di una presa di turno, così come non c'è restrizione al numero dei partecipanti ad una conversazione. A questo proposito Levinson scrive:

> […] inoltre, data la flessibilità delle unità sintattiche costitutive dei turni e i modi di proseguire consentiti dalla regola (1c)[7], non ci sono limiti rigidi per l'estensione temporale del turno; infine, nessun partecipante è escluso a priori, il numero dei partecipanti in una conversazione può variare (Levinson, 1993: 303-4).

L'altra serie di restrizioni riguarda la realizzazione delle *azioni comunicative possibili*, cioè cosa si deve e si può fare quando si ha la

[7] La regola (1c) si riferisce all'insieme di regole che compongono il meccanismo di allocazione dei turni di parola nella conversazione quotidiana. La descrizione di questo sistema è stata presentata nel famoso Sacks, Schegloff and Jefferson (1974).

parola. Tali restrizioni intervengono nella *costituzione di un tipo di sequenza adeguato* per l'adempimento dello scambio comunicativo. Gli scambi poetici sono co-costruiti attraverso dei tipi specifici di coppie adiacenti investite da restrizione nell'organizzazione locale del passaggio tra i turni. Esistono vari formati di costruzione dello scambio che differenziano l'improvvisazione dei *saluti* e dei *congedi* dall'improvvisazione dei *contrasti*. I formati di costruzione dei turni non sono intercambiabili tra loro. Infine, i vincoli sul tipo di coppie adiacenti possibili nello scambio improvvisato, determinano una restrizione delle scelte nel *sistema di preferenze* per la costituzione delle coppie che in questo modo orienta il *formato di produzione* e le sue sotto sequenze verso un tipo di efficacia pragmatica precisa. Le restrizioni agiscono quindi anche sul valore pragmatico-azionale della presa di turno mono- o multi-azionale che sia. Tale effetto sembra determinato dalla coincidenza tra *collocazione dei turni di parola* e *posizione funzionale dei turni*. Levinson introduce questa distinzione nel suo studio dedicato alle pre-sequenze (1993: 350): «Occorre introdurre una distinzione tra localizzazione e posizione funzionale del turno» ed afferma che nella conversazione quotidiana

> *l*e situazioni possono cambiare ma in genere ogni posizione assume determinate forme, indipendentemente dalla collocazione assoluta all'interno di una sequenza di turni (*ibidem*).

Secondo Levinson, dunque, nella conversazione ordinaria non c'è coincidenza tra posizione e funzione del turno. Per contro, nell'improvvisazione poetica, la posizione dei turni o delle UCT (Unità di Costitutiva del Turno) corrisponde di solito ad un loro valore azionale, ha quindi uno statuto funzionale rispetto ad una sequenza data. Ricordo, ad esempio, la posizione dell'atto performativo esplicito di saluto enunciato nel distico finale di strofa nelle ottave di apertura delle competizioni e le posizioni delle azioni protocollari nelle ottave a contrasto (vedi 6.2.1). L'applicazione delle restrizioni nel formato di produzione orienta la coordinazione dello scambio verso un ordinamento preciso dello scambio con cui si costruisce uno specifico valore pragmatico-azionale. In tal modo esercita una forma di controllo sull'interazione in particolare nello scambio oppositivo.

8 Conclusioni

L'insieme dei protocolli e dei vincoli che ho presentato sono gli elementi che distinguono le *performance* poetiche dagli scambi della conversazione ordinaria. Essi intervengono sulle molteplici dimensioni dell'attività dialogica co-costruita nelle improvvisazioni in ottava rima ed impongono agli improvvisatori di coordinare la loro attività secondo alcuni principi operativi che si ancorano alle risorse formali rese disponibili dall'articolazione intermodale della forma dell'espressione. Un principio di *equità e stabilizzazione* relativo alla transazione e alla durata del turno di parola, riferibile alla versificazione isometrica. Un principio di *reciprocità* nelle opportunità dell'azione relativo alla pianificazione pragmatica, riferibile alla articolazione tra melodia e metrica. Un principio di *accordo* (*parziale*) con l'interlocutore, relativo all'espressione del consenso e del dissenso, riferibile alla articolazione dei moduli melodico-metrici con il sistema di rime.
Questi principi, indispensabili per la buona esecuzione delle improvvisazioni, sono iscritti nella *prassi abituale* condivisa dai membri. Sono cioè un sapere procedurale e non dichiarativo, incarnato nella *tecnica* esperta degli improvvisatori che va aldilà dei contenuti circostanziali delle singole *performance*. In questo senso saper improvvisare bene in ottava rima significa possedere una forma di *saggezza dialogica* incorporata lentamente attraverso l'acquisizione di una competenza socio-cognitiva specifica rispetto a quella della comunicazione parlata ordinaria. Significa la condivisione di un *habitus espressivo* che rivela la sedimentazione di un'*etica del dialogo*. Con le sue proprietà sistemiche l'improvvisazione dialogica toscana abolisce la gerarchia tra gli interlocutori e si fa garante di una giurisprudenza per la gestione del diritto di replica nell'interazione oppositiva. La natura *deontica* della sua poetica codifica un ordinamento dell'esecuzione che impone una *solennità* necessaria alla stabilità dell'interazione nei casi di relazioni sociali che possono presentare elementi di rischio o di pericolo per l'ordine pubblico. Il mio lavoro suggerisce a questo proposito un possibile collegamento tra la tradizione culturale dell'improvvisazione in ottava rima e l'attività socio-economica della transumanza praticata dalle comunità dei pastori dell'Appennino tosco-emiliano verso la Maremma grossetana. La transumanza è sempre stata un'occasione di incontri e

di scontri reiterati nel tempo tra la popolazioni degli agricoltori stanziali ed i "forestieri" transumanti per la gestione di una prossimità nello spazio geografico. I contatti tra queste differenti comunità umane sono avvenuti secondo modalità consuetudinarie plurisecolari fino ai primi decenni del XX secolo. Secondo alcune ricostruzioni storiche relative alla transumanza interna alla regione Toscana (Massaini, 2005) i pastori dovevano negoziare con le comunità locali il transito o la permanenza transitoria – per esempio il pernottamento – nei terreni agricoli di questi ultimi oppure ridefinire frequentemente i confini dei pascoli. Questo avveniva sia durante il percorso della transumanza lungo i tratturi – compiuto due volte all'anno dai pastori per spostare i loro greggi dall'Appennino verso la Maremma e viceversa – sia durante i soggiorni stagionali nei terreni di pascolo presi in affitto dalle comunità dei pastori dal mese di settembre al mese di maggio. In genere il *vergaio*, il capo della comunità transumante, aveva la responsabilità di questa delicata negoziazione che era di fondamentale importanza per la vita quotidiana della comunità ed al contempo presentava spesso notevoli rischi d'insuccesso. È importante ricordare brevemente che la gestione dei diritti di pascolo e di tutela dei diritti/doveri dei pastori transumanti ha portato la repubblica di Siena fin dal 1300 alla produzione di una amministrazione specifica dei Paschi della Maremma con la creazione di un vero e proprio statuto della Dogana dei Paschi Maremmani emanato nel 1419 (Massaini, 2005:46). Con la nascita di questa istituzione la repubblica aveva la completa competenza nella distribuzione e riscossione delle concessioni del pascolo pubblico[8] date ai vergai o ai proprietari di animali.

Confrontando le possibili interazioni tra stanziali e transumanti – stabilizzate da pratiche consuetudinarie millenarie, anche comunicative – e la specializzazione dei quadri di attività dialogica delle competizioni poetiche della Maremma, si scoprono delle straordinarie convergenze che vale la pena di approfondire ulteriormente.

[8] La regolamentazione in termini economici dei pascoli maremmani consentirà l'istituzione, nel 1472, del primo istituto di credito fondiario, quello che in seguito sarà chiamato Monte dei Paschi di Siena. L'imposta di Dogana dei Paschi sarà abolita soltanto nel 1778 dal Granduca Pietro Leopoldo di Lorena (cfr. Massaini, 2005).

La mia ipotesi interpretativa suggerisce l'associazione tra questi due campi di attività e propone di considerare una gara d'improvvisazione poetica toscana – così estranea alle giurie degli esperti – come la rappresentazione pubblica di un elaborato sistema simbolico che indicizza una realtà esperienziale socio-storica ed esprime una concezione culturale specifica sulle forme di gestione della prossimità. In tal senso una competizione o riunione tradizionale d'improvvisazione in ottava rima, sarebbe un cerimoniale, presumibilmente sedimentatosi nel tempo all'interno di un'area geo-culturale, composto da più "rituali di parola" con il quale si trasmette, ancora oggi, una precisa *memoria sociale* relativa alla costruzione delle relazioni con il "forestiero". Una tale ipotesi, del tutto trascurata nei lavori che mi hanno preceduto, invita a riesaminare *la natura* di questa antica tradizione orale che sembra rivelare degli aspetti di una sorprendente attualità sociale e politica.

Riferimenti bibliografici

Agamennone, M. (1999) Maestri della voce, maestri del tempo, in P. Nardini (ed.) *L'Arte del dire. Atti del Convegno di studi sull'improvvisazione poetica*, Biblioteca Chelliana, Archivio delle Tradizioni Popolari della Maremma Grossetana, pp.139-61.

– (2002) Modi del contrasto in ottava rima, in M. Agamennone and F. Giannattasio (eds) *Sul verso cantato. La poesia orale in una prospettiva etnomusicologica*, Padova: Il Poligrafo, pp.163-223.

Austin, J.L. (1962) *How to do Things with Words.* Oxford: Oxford University Press.

Beltrami, P. (1994) *La metrica italiana*. Bologna: il Mulino.

Fornel, M. de (1989) Rituel et sens du rituel dans les échanges conversationnels, in *Le parler frais d'Erving Goffman*, Paris: Éditions de Minuit.

Fornel M. de. and G. Tiezzi (2008) Approccio sequenziale al duello verbale, in M. Pettorino, A. Giannini, M. Vallone and R. Savy (eds) *Atti di La Comunicazione Parlata 2006*, Vol. I. Napoli: Liguori, pp. 1066-1089.

Goffman, E. (1986) 'La condition de félicité'. *Actes de la recherche en sciences sociales*. Vol. 64/1: 63-78, Vol. 65/1: 87-98.

Levinson, S.C. (1983) *Pragmatics*. Cambridge: Cambridge University Press [trad. it. (1993) *La Pragmatica*. Bologna: il Mulino]

Massaini, M. (2005) *Transumanza. Dal Casentino alla Maremma. Storie di uomini e di armenti lungo le antiche dogane*. Roma: Aldo Sara Editore.

Sakcs, H., E. Schegloff and G. Jafferson (1974) 'A simplest systematic for the organisation of turn-taking for conversation'. *Language* 50, pp. 696-735.

Silverstain, M. (1998) 'Contemporary Transformations of local Linguistic Communities'. *Annual Review of anthropology* vol. 27, pp.401-26.

Tiezzi, G. (2008) Le ragioni della rima. Studio sull'improvvisazione di un'ottava di saluto. *Documenti di lavoro e pre-pubblicazioni*, n. 376/2008 – serie D, ed. Centro Internazionale di Semiotica e di Linguistica, Università di Urbino.

Zumthor, P. (1984) *Introduction à la poésie orale*. Paris: Seuil [trad. it. (2001) *La presenza della voce. Introduzione alla poesia orale.* Bologna: il Mulino).

PARTE VI

COMUNICAZIONE PARLATA E PATOLOGIE DEL LINGUAGGIO

Ipotesi per una valutazione dei deficit afasici all'interfaccia tra sintassi e scelte pragmatico-discorsive

Camilla Bettoni, Maria Elena Favilla°, Lucia Ferroni^*
*Università di Verona, °Università di Modena e Reggio Emilia,
^Università di Pisa e Az. USL di Lucca
camilla.bettoni@univr.it; *elena.favilla@unimore.it*; *l.ferroni@usl2.toscana.it*

Abstract

In aphasic studies, the current modular approach has proven useful to single out, assess, and treat different components involved in the processing of words. Its extension beyond the word is however less uncontroversial. While several defective syntactic and pragmatic skills have been isolated in agrammatic patients, their co-occurrence seems more probabilistic than regular. On the other hand, studies in L2 acquisition have identified universal developmental sequences. Processability Theory (Pienemann 1998; Pienamann, Di Biase and Kawaguchi, 2005) in particular can explain the implicational nature of some grammatical sequences on the basis of the processing cost required during speech production. In this paper our focus is on Italian, a language with a syntax highly sensitive to pragmatic and discourse choices. We test 6 agrammatic patients with no relevant phonological or lexical impairments with the aim of establishing whether the implicational nature of morpho-syntactic and discourse-pragmatic development hypothesised for L2 learners holds true, in the opposite direction, also for agrammatic L1 speakers. Our results appear to support our hypothesis, promising useful clinical applications for both assessing and treating aphasic deficits.

1 Afasia e Teoria della Processabilità

Il linguaggio afasico fornisce dati utili a comprendere il funzionamento dei meccanismi coinvolti nell'elaborazione linguistica. I diversi modi in cui le singole abilità linguistiche e i singoli livelli di analisi del linguaggio possono essere conservati o danneggiati a seguito di una lesione cerebrale permettono di raccogliere informazioni e di verificare o formulare ipotesi sul funzionamento del linguaggio normale. La possibilità di correlare la capacità di svolgere determinati compiti linguistici con specifiche aree cerebrali fornisce, inoltre, indicazioni sulle aree cerebrali coinvolte nello svolgimento di quei compiti linguistici. A loro volta, i risultati ottenuti dalla

linguistica per le diverse unità di analisi del linguaggio e le diverse strutture linguistiche costituiscono un'utile risorsa per comprendere e curare nel modo più efficace e rapido possibile i disturbi del linguaggio. In questo senso un'interazione fra le discipline che individuano il proprio oggetto principale nel funzionamento del linguaggio normale, ed in particolare la psicolinguistica, e le discipline che privilegiano lo studio del linguaggio patologico, come la neurologia, può rivelarsi reciprocamente fruttuosa, se non addirittura necessaria.

Più specificamente, in ambito neuropsicologico, l'approccio cognitivo modulare offre un modello in grado di rendere conto dell'elaborazione semantico-lessicale di singole parole e di individuare le diverse componenti che vi contribuiscono (Patterson, 1986; Caramazza, 1988, 1997; 2000; Caramazza and Hillis, 1990; Hillis and Caramazza, 1991; 1995; Miceli, Capasso and Caramazza, 1994; Miceli, Benvegnù and Capasso, 1997; Miceli and Capasso, 1997). Questo modello sul funzionamento normale del linguaggio consente di mettere a punto strumenti di valutazione volti ad individuare le componenti danneggiate nei soggetti afasici e di impostare trattamenti riabilitativi specifici per ciascuna componente. L'utilità di tale approccio ha indotto ad estenderlo ai decifit morfo-sintattici e di interfaccia tra sintassi e pragmatica (per es., Schwarz, Fink and Saffran, 1995). Rispetto al modello semantico-lessicale, tuttavia, questa estensione oltre il livello della parola appare più controversa dal punto di vista scientifico e meno soddisfacente sul piano clinico. Negli ultimi anni sono stati compiuti diversi tentativi di usare teorie sintattiche per spiegare la grammatica degli afasici (per es., Caplan and Hildebrandt, 1988; Caramazza and Miceli, 1991; Grodzinsky, 1986; 2000; Friedman, 1997; 2002; Luzzatti and De Bleser, 1996). Per quanto utili sul piano clinico per trattare singoli aspetti della produzione, i metodi elaborati non forniscono un modello generale per valutare i molteplici aspetti della competenza sintattica e individuare in modo sistematico le componenti da riabilitare (per una presentazione di questi metodi si rimanda a Ferroni, 2009).

In ambito linguistico, gli studi sull'acquisizione di lingue seconde propongono oggi alcune sequenze universali per lo sviluppo delle strutture grammaticali. Quando la loro natura è implicazionale esse potrebbero offrire un contributo non trascurabile alla soluzione del problema della grammatica afasica, sul piano sia teorico sia

metodologico fornendo la possibilità di ritrovare nella dissoluzione delle varie strutture tra i pazienti afasici le gerarchie inverse a quelle della loro comparsa tra gli apprendenti. Se così fosse, questa applicazione comporterebbe dei vantaggi anche sul versante della linguistica acquisizionale, perché risulterebbe un'importante conferma per le teorie elaborate.

Una teoria che si presta in modo particolare ad un'interazione tra linguistica acquisizionale e studio dei disturbi del linguaggio è la Teoria della Processabilità (TP), elaborata da Manfred Pienemann (1998; Pienemann *et al*, 2005; e per un quadro recente e unitario della teoria, Bettoni and Di Biase, in press; Di Biase and Kawaguchi, in press). La TP è una teoria dell'apprendimento grammaticale di una L2 fondata su due solidi pilastri teorici: il modello psicolinguistico di Levelt (1989) per quanto riguarda la sequenza temporale delle fasi di produzione dell'enunciato, e la Grammatica Lessico-Funzionale di Bresnan (2001) per quanto riguarda la descrizione della lingua target. Secondo la TP gli stadi universali dello sviluppo linguistico si spiegano in base al costo dell'elaborazione cognitiva richiesta dalla produzione delle strutture grammaticali. L'individuazione di queste procedure di elaborazione che determinano gli stadi permette di predire le sequenze implicazionali in cui vengono apprese le strutture specifiche di singole lingue. Formulata inizialmente su dati soprattutto inglesi e tedeschi, la TP è adesso verificata non solo su lingue tipologicamente distanti, come il giapponese e l'arabo, ma anche in situazioni differenti, come quelle del bilinguismo precoce e dello *Specific Language Impairment*. Per l'italiano L2, le sequenze di sviluppo sono esposte in Di Biase and Kawaguchi (2002); Di Biase (2007); Bettoni, Biase and Ferraris (2008); Bettoni, Di Biase and Nuzzo (2009); Bettoni and Ginelli (in press).

In questo contributo presentiamo i primi risultati di una ricerca volta ad esplorare l'applicabilità delle ipotesi proposte dalla TP per l'apprendimento alle condizioni di perdita afasica. Per raccogliere i dati, è stato messo alla prova un protocollo di rilevazione del parlato volto ad elicitare alcune strutture grammaticali che rientrano nelle ipotesi della TP per valutare se le implicazioni riscontrate per gli apprendenti della L2 valgono anche per i parlanti agrammatici di italiano L1. In particolare, viene valutato il ruolo svolto dalla natura delle corrispondenze che intercorrono tra le gerarchie delle tre strutture parallele che secondo Bresnan (2001) rappresentano

formalmente la frase: la struttura argomentale (prima l'agente, poi il beneficiario, l'esperiente ecc., fino al locativo), quella funzionale (prima il soggetto, poi l'oggetto diretto ecc.) e quella per costituenti (prima posizione, seconda ecc.). A questo scopo, viene analizzata la produzione di frasi contenenti verbi passivi, topicalizzazioni dell'oggetto e verbi del tipo che Pinker (1984) definisce eccezionali.

2 Le ipotesi

Secondo l'Ipotesi dell'Allineamento Non Marcato della TP, nell'interlingua, la sintassi di una frase semplice, dichiarativa, affermativa e pragmaticamente neutra si sviluppa partendo dalla corrispondenza non marcata tra gli elementi della struttura argomentale (i ruoli tematici) e quelli della struttura costituente (i costituenti). In questo modo la frase risulta organizzata secondo l'ordine canonico, e i singoli costituenti possono rimanere aspecifici riguardo alle loro funzioni sintattiche senza creare problemi di fraintendimento per l'ascoltatore, poiché in prima posizione c'è l'agente, in seconda il beneficiario e così via. Tuttavia, quando per motivi pragmatico-discorsivi il parlante intende assegnare diversa prominenza topicale o focale ai costituenti (Levelt, 1989), e quindi in italiano scompaginarne l'ordine canonico, deve poter essere in grado di marcare le funzioni sintattiche.

Questo implica una corrispondenza marcata tra gli elementi gerarchici dei tre livelli della frase, che può realizzarsi in due modi. Quando è la struttura argomentale che non è mappata in modo canonico su quella funzionale, la TP propone l'Ipotesi del Mapping Lessicale. Ne costituiscono un esempio i verbi passivi, che per la loro natura lessicale richiedono che l'agente non sia soggetto, bensì complemento indiretto, quando non venga persino soppresso (per es. *Pierino è stato punto da una zanzara*). Quando invece è la struttura costituente che si allinea in modo marcato su quella funzionale, la TP propone l'Ipotesi del Topic. Ne costituisce un esempio la topicalizzazione dell'oggetto, in cui l'elemento in prima posizione non è soggetto, bensì oggetto (per es. *Pierino l'ha punto una zanzara*).

La progressione dell'interlingua prevede che gli apprendenti imparino ad elaborare prima le frasi nelle quali le gerarchie delle tre strutture si allineano in modo non marcato, e che solo in un secondo tempo riescano ad elaborare frasi con allineamenti marcati. Ciò significa, per

esempio, che in una determinata fase dello sviluppo un apprendente sarà in grado di produrre in modo grammaticalmente corretto le frasi nelle quali il soggetto coincide con l'agente in prima posizione, mentre commetterà errori nel produrre frasi in cui il soggetto non coincide con l'agente o non si trova in prima posizione.
Questi due tipi di corrispondenza marcata si possono trovare insieme con i verbi che Pinker (1984) definisce eccezionali, come *piacere* e *mancare*. In questo caso le difficoltà aumentano, poiché l'apprendente deve marcare sia in prima posizione l'esperiente con funzione dativa, sia in posizione postverbale il tema con funzione di soggetto (per es. *a Pierino serve l'ombrello*).
Tra queste tre costruzioni, ipotizziamo una sequenza che prevede in un primo momento l'apprendimento dei verbi passivi e della topicalizzazione dell'oggetto, e poi dei verbi eccezionali. Fondiamo questa ipotesi sul fatto che i verbi passivi mantengono l'ordine canonico dei costituenti (SVO), e la topicalizzazione mantiene inalterate le funzioni sintattiche dei ruoli tematici, mentre i verbi eccezionali prevedono, come appena detto, tutti e due i tipi di corrispondenze marcate.
Se le sequenze evolutive della TP valgono anche nella direzione inversa per gli afasici, dall'analisi della produzione linguistica dei soggetti afasici dovrebbero emergere le stesse sequenze, e cioè dovremmo trovare che i pazienti sanno produrre (a) frasi corrette con l'allineamento non marcato e frasi scorrette con quello marcato; (b) frasi più corrette con i verbi passivi e con la topicalizzazione dell'oggetto, e meno corrette con i verbi eccezionali. Poiché in italiano l'ordine dei costituenti è flessibile, gli errori grammaticali che conseguono alla mancata assegnazione funzionale dei ruoli tematici e dei costituenti saranno di natura morfologica.

3 Il protocollo

Al fine di raccogliere dati sulle capacità di elaborazione delle tre strutture indagate è stata adattata alle specifiche esigenze e condizioni dei pazienti afasici parte del materiale del protocollo elaborato da Nuzzo *et al* (in press) per elicitare la produzione di strutture facoltative per l'analisi dell'interlingua italiana. La capacità di utilizzare verbi passivi, topicalizzazioni dell'oggetto e verbi

eccezionali è infatti difficilmente valutabile su dati di produzione spontanea, dove queste strutture non sono necessariamente frequenti: ad esempio, nell'analisi della conversazione di un adulto italiano sano, Bettoni, Di Biase and Nuzzo (2009) riscontrano soltanto 4 casi di topicalizzazioni dell'oggetto su 255 frasi contenenti un verbo lessicale. I task utilizzati (v. tab. 1) includono i seguenti:

1) descrizione di brevi filmati per elicitare l'uso di verbi passivi. Vengono mostrati filmati della durata di pochi secondi. Al soggetto viene chiesto di dire cosa succede iniziando la risposta con la parola che appare sullo schermo all'inizio di ogni filmato, che viene anche letta ad alta voce dal ricercatore (per es., *il leone viene morso dal serpente*);
2) picnic degli animali per elicitare la topicalizzazione. In una prima fase, si spiega al soggetto che alcuni animali organizzano un picnic e che ogni partecipante porta qualcosa da mangiare. Il ricercatore mostra di volta in volta due cartoncini, l'uno raffigurante un tipo di cibo (per es., *i panini*), l'altro un animale (per es., *il cane*). Al soggetto viene chiesto di dire chi porta che cosa, iniziando dal cartoncino che viene mostrato per primo, così da produrre frasi come *i panini li porta il cane* o *il cane porta i panini*. In una seconda fase, si racconta che è scoppiato un temporale e che gli animali hanno messo in salvo le loro vivande. Di fronte agli stessi cartoncini in abbinamenti diversi, il soggetto deve dire chi ha salvato cosa, procedendo come sopra. Le due fasi sono dunque concepite per elicitare topicalizzazioni prima al presente e poi al passato prossimo, così da poter valutare anche l'accordo del participio;
3) descrizione di immagini per elicitare i verbi eccezionali. Vengono mostrati tre cartoncini, due raffiguranti immagini e uno sul quale è scritta la forma infinita di un verbo eccezionale. Il soggetto deve costruire una frase di senso compiuto seguendo l'ordine degli elementi raffigurati sui cartoncini.

Task	**Strutture attese**	**Contesti proposti**
descrizione di brevi filmati	verbi passivi	7
picnic degli animali	topicalizzazione dell'oggetto	8
descrizione di immagini	verbi eccezionali	7

Tabella 1: I task utilizzati per la raccolta dei dati

A questi tre task specifici, si è aggiunto il racconto di un brano di circa dieci minuti tratto dal film *Modern Times* di Chaplin, permettendo così al soggetto di esprimersi più liberamente, e a noi di verificare che,

in frasi con corrispondenze non marcate, tutti possedessero solide le strutture morfologiche necessarie alla costruzione delle frasi target con corrispondenze marcate.

4 I soggetti afasici

Le lesioni che possono dare luogo ad afasia sono di varia natura e sono determinate da ictus ischemici o emorragici, ma anche da traumi cranici, vari tipi di patologie cerebrali o anche conseguenza di interventi chirurgici. Le lesioni ischemiche e le emorragiche sono quelle che meglio si prestano a studi sul funzionamento del linguaggio. Infatti, essendo di solito più circoscritte, determinano disturbi molto specifici e per lo più limitati alle funzioni linguistiche, lasciando integre le altre abilità cognitive (memoria, attenzione ecc.), che comunque condizionano le capacità di elaborazione linguistica. Per illustrare gli ulteriori criteri di selezione dei pazienti è opportuno soffermarsi brevemente su alcune questioni relative alla valutazione dei pazienti afasici così come viene condotta nella pratica clinica. Ci limiteremo qui ad alcune indicazioni generali, mentre, per una trattazione più approfondita, si rimanda a Favilla (2003), Favilla and Ferroni (2009).

Nella prospettiva clinica dell'inquadramento del paziente, i diversi tipi di afasia vengono di solito classificati in base alle capacità linguistiche del paziente e alla sede della lesione. Relativamente al primo aspetto, in una prima fase vengono valutate le capacità di produzione, comprensione e ripetizione, che, come noto, possono essere danneggiate in misura diversa. Della produzione si valuta innanzitutto l'eloquio, che viene definito fluente o ridotto in relazione al numero di parole al minuto, alle capacità prosodiche e alla presenza di elementi lessicali appartenenti alle varie classi grammaticali. Schematizzando una situazione ben più complessa ed articolata, i pazienti con eloquio ridotto producono poche parole per minuto, quasi esclusivamente parole singole, costituite per lo più da nomi nella forma non marcata, con un andamento prosodico alterato. Viceversa, i pazienti con eloquio fluente tendono a produrre un numero elevato di parole per minuto, appartenenti a tutte le classi (nomi, verbi e aggettivi nelle varie forme flesse, parole funzionali ecc.), combinate tra loro in modo non controllato e unite a parole possibili ma inesistenti, così che il risultato è una produzione caratterizzata da un andamento prosodico

normale, apparentemente corretta ma spesso priva di senso. Di solito, l'eloquio ridotto è associato a lesioni anteriori, più precisamente del lobo frontale dell'emisfero sinistro (area di Broca), mentre l'eloquio fluente è associato a lesioni posteriori, del lobo temporale (area di Wernicke). La comprensione risulta più spesso compromessa in caso di lesione posteriore, mentre la ripetizione può essere compromessa in caso di lesione sia anteriore sia posteriore, ma anche come conseguenza di una lesione che interrompe le connessioni tra aree del linguaggio posteriori e anteriori.

Queste correlazioni tra sede della lesione e comportamento linguistico sono da intendersi in senso generale come conseguenza del fatto che le varie abilità cognitive, tra cui quelle linguistiche, richiedono l'attivazione di determinati circuiti cerebrali che interessano specifiche aree cerebrali (per una discussione sui rapporti tra localizzazioni cerebrali e deficit neuropsicologici, vedi Favilla, Ferroni and Giannini, 2009: 129-34). È importante considerare, tuttavia, che si possono riscontrare anche dei casi nei quali tra sede della lesione e tipo di afasia non sono individuabili correlazioni tipiche, così che, ad esempio, è possibile trovare un paziente con lesione posteriore e manifestazioni cliniche dell'afasia di Broca.

Al di là della classificazione anatomo-clinica, inoltre, tutti i pazienti afasici possono presentare difficoltà di elaborazione morfosintattica, anche se con manifestazioni concrete molto diverse tra loro. Come si è anticipato, i disturbi di elaborazione morfosintattica e l'agrammatismo in generale non sono ancora del tutto chiari. In particolare non è chiaro se le differenze nel tipo di comportamento linguistico possano essere almeno in parte ricondotte al tipo di afasia.

Questi elementi hanno guidato la scelta dei nostri pazienti. Le prove descritte in § 3 sono state somministrate a sei pazienti di età compresa tra 52 e 79 anni e scolarità compresa tra 5 e 11 anni. Sono stati scelti solo pazienti con lesioni ischemiche o emorragiche e, relativamente al tipo di afasia, l'unico requisito per l'inclusione è stato la mancanza di disturbi fonologici e lessicali gravi, che avrebbero reso impossibile la raccolta delle strutture target. Come si può vedere da tab. 2, tre pazienti sono affetti da afasia di Broca, un paziente da afasia originariamente di Broca poi evoluta in afasia transcorticale motoria (con disturbo nella produzione, ma, diversamente dall'afasia di Broca, non nella ripetizione), e due pazienti con afasia di Wernicke, una delle

quali evoluta in afasia di conduzione (eloquio fluente, comprensione integra e difficoltà nella ripetizione).

Paziente	Variabili socioculturali	Natura e sede della lesione	Tipo di afasia
Sandra	52 anni, scolarità 8 anni, commessa in pensione	ictus emorragico, frontale sinistra	Broca
Ida	59 anni, scolarità 5 anni, operaia in pensione	ictus emorragico, fronto-temporo-parietale, cortico sottocorticale sinistra	Broca
Fulvia	66 anni, scolarità 8 anni, casalinga	ictus emorragico, fronto-temporale sinistra	Broca
Luigi	69 anni, scolarità 8 anni, cameriere in pensione	ictus ischemico, parietale sinistra	Broca → transcorticale motoria
Anna	66 anni, scolarità 11 anni, impiegata in pensione	ictus ischemico, occipitale sinistra	Wernicke → conduzione
Giulia	79 anni, scolarità 5 anni, operaia in pensione	ictus ischemico, occipitale sinistra + temporo-fronto-parietale sinistra	Wernicke

Tabella 2: I pazienti

Tutti i pazienti sono stati valutati con un test standardizzato, l'Esame del linguaggio-II elaborato da Ciurli, Marangolo and Basso (1996), che permette di raccogliere informazioni sulle abilità linguistiche nei vari compiti (produzione ripetizione, comprensione e copia di parole, non parole e frasi, nelle modalità sia orale sia scritta). I risultati, riportati in tab. 3, mostrano che gli strumenti esistenti per la valutazione del linguaggio afasico classificano i disturbi di questi pazienti come molto lievi – almeno secondo questo test, costituito da prove che richiedono l'elaborazione di unità lessicali considerate singolarmente o di frasi isolate caratterizzate da una sintassi semplice.

	Produzione orale			Comprens. orale			Ripetizione			Scrittura			Lettura (comprens.)			Lettura ad alta voce			Dettato		
Pazienti	descrizione immagine	denominazione nomi	denominazione verbi	parole	parole semanticam. affini	frasi	parole	non parole	frasi	descrizione immagine	denominazione nomi	denominazione verbi	parole	parole semanticam. affini	frasi	parole	non parole	frasi	parole	non parole	frasi
S	90	100	97	100	100	100	97	90	100	90	90	95	100	100	90	100	90	100	100	90	100
F	60	100	80	100	100	100	100	85	100	70	100	70	100	100	100	100	90	100	85	100	95
I	80	90	80	100	100	100	95	80	70	90	95	100	100	100	100	97	75	95	97	80	90
L	95	95	80	100	100	100	100	100	100	95	90	90	100	100	100	100	100	95	90	85	70
A	80	100	100	100	100	90	100	85	80	100	100	100	100	100	100	100	97	90	100	95	100
G	90	95	80	100	100	100	100	75	85	60	100	60	100	100	70	100	90	97	92	90	75

Tabella 3: Le prestazioni dei pazienti secondo un test solitamente usato in ambito afasico (≥90%: normale; 80-90%: borderline; <80%: deficitaria)

Nonostante i risultati quasi nella norma ottenuti in queste prove, i sei pazienti presentano difficoltà linguistiche di elaborazione sintattica che non consentono lo svolgimento di una vita sociale e lavorativa normale. Disporre, quindi, di strumenti di valutazione e di teorie in grado di rendere conto in modo più sensibile dei disturbi relativi all'elaborazione grammaticale sarebbe essenziale per impostare trattamenti riabilitativi mirati ed efficaci. Le strutture target della nostra indagine costituiscono, come evidenziato da Levelt (1989), utili alternative all'ordine canonico delle parole e al mapping non marcato. Ciò le rende particolarmente importanti nelle interazioni quotidiane perché consentono di richiamare e mantenere l'attenzione dell'interlocutore. Inoltre, per quei soggetti che, come gli afasici, presentano difficoltà linguistiche, può risultare preziosa qualsiasi strategia che renda la produzione più immediata ed efficace. In questo senso, un lavoro mirato a queste strutture renderebbe disponibile una maggiore gamma di possibilità cui attingere per far fronte alle difficoltà di elaborazione lessicale che spesso caratterizzano la produzione afasica e che possono rendere impaziente l'ascoltatore.

5 Risultati

Per tutti i sei pazienti, i dati elicitati mostrano come in casi di allineamento non marcato la produzione avvenga senza errori di tipo morfologico. Su questa base, se si trattasse di apprendenti di L2 anziché di pazienti afasici valutati in L1, potremmo collocare tutti i soggetti nello stadio evolutivo precedente all'emergenza di strutture contenenti verbi passivi, topicalizzazioni dell'oggetto e verbi eccezionali. In particolare, essi sanno accuratamente coniugare le persone del verbo, concordare il participio passato dei verbi con ausiliare *essere* e flettere i clitici, e dunque possiedono i mezzi per marcare la funzione sintattica dei ruoli tematici e dei costituenti in caso di allineamento marcato. Il farlo effettivamente costituisce un ulteriore passo in avanti.

Per quanto riguarda le nostre tre strutture target, per valutare le capacità dei soggetti nella loro elaborazione è stato calcolato il numero di casi di produzione corretta rispetto al totale di contesti nei quali ciascun paziente ha cercato di produrre una determinata struttura. Non sono stati considerati i casi in cui è stata prodotta una frase con una struttura diversa da quella target, perché quando le strutture target sono facoltative, la mancata produzione non implica necessariamente l'incapacità di usarle.

Confrontando nei diversi pazienti la produzione delle tre strutture considerate complessivamente, emerge una prestazione differenziata fra i soggetti, che oscillano tra il 70% di produzioni corrette di Sandra e il 10% di Ida (v. fig. 1). Con riferimento alle singole strutture, considerando globalmente il gruppo dei soggetti afasici, l'ipotesi di partenza risulta confermata: le frasi con verbo passivo e con topicalizzazione sono prodotte con un numero di errori pressoché uguale, e minore rispetto a quelle con un verbo eccezionale (v. fig. 2).

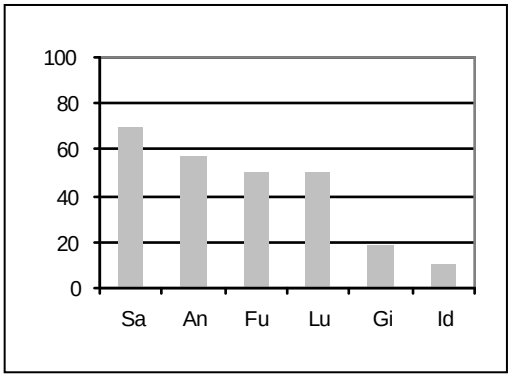


Figura 1: Le tre strutture – produzione corretta (%)

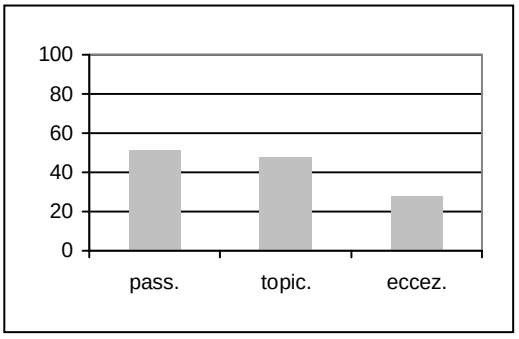


Figura 2: I sei pazienti – produzione corretta (%)

Se si considerano invece le prestazioni dei singoli pazienti relativamente alle singole strutture il quadro risulta meno chiaro (v. fig. 3). Infatti, rispetto ai verbi eccezionali, Sandra ha maggiori difficoltà con le topicalizzazioni, e Anna con i verbi passivi.

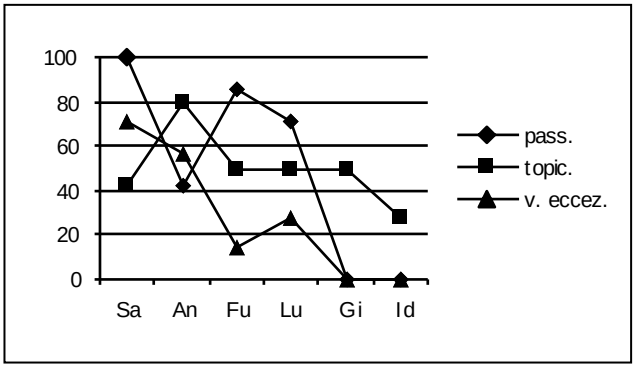


Figura 3: Le tre strutture nei sei pazienti – produzione corretta (%)

Ulteriori indicazioni interessanti possono essere ricavate dall'analisi qualitativa delle produzioni dei pazienti relativamente ai tre tipi di strutture.
Per quanto riguarda i verbi passivi, quando non è stata prodotta la frase target, in alcuni casi i pazienti ignorano la consegna e iniziano la descrizione con un referente diverso da quello proposto; ad esempio, nel filmato che mostra una donna che viene scippata, nonostante la richiesta di iniziare la frase con *la donna* viene detto **questo ragazzo gli scippa alla signora*. In altri casi, invece, i pazienti iniziano la frase con l'elemento richiesto, ma utilizzando il verbo attivo e cambiando così il contenuto proposizionale, come nella frase **il leone ha pinzato la serpe* quando si chiedeva *il leone è morso dal serpente*. In altri casi ancora, più rari, viene prodotta una frase priva di senso, come **il maialino spacca il salvadanaio* invece che *il maialino/un salvadanaio viene distrutto*. Il prodotto comunque è sempre una frase in cui l'agente in prima posizione ha la funzione di soggetto, confermando la problematicità di utilizzare un mapping diverso da quello canonico nella corrispondenza tra la struttura argomentale e quella funzionale. Inoltre è importante sottolineare che nessuna di queste frasi non target presenta errori morfologici.
Per quanto riguarda la topicalizzazione dell'oggetto, invece, mentre non presenta alcun problema l'ordine non canonico delle parole, subentrano difficoltà nei tre accordi. L'accordo tra verbo e soggetto è corretto nel 90% dei casi se il soggetto è non marcato (singolare), come nella frase *le ciliegie le porta il cane*, mentre è corretto solo nel 16% dei casi se il soggetto è marcato (plurale), come nella frase *l'insalata la *porta i gatti*. Il sintagma topicalizzato sembra creare meno problemi di accordo con il clitico; comunque, ancora una volta il numero di frasi corrette è maggiore quando il sintagma topicalizzato è morfologicamente non marcato (maschile singolare), come *il salame lo porta le galline*, che registrano 88% di usi corretti, rispetto a quando è marcato (femminile e/o plurale), come *l'insalata *lo porta i gatti*, che registrano 73% di usi corretti. Il maggior numero di errori viene prodotto, prevedibilmente, quando sono marcati sia il soggetto sia l'oggetto topicalizzato. Infine, nelle frasi al passato, l'accordo del participio con l'oggetto è sempre prodotto correttamente quando l'ausiliare è preceduto dal clitico (nell'unico caso in cui il participio non è concordato con l'oggetto, l'oggetto è plurale e, dunque,

marcato, mentre il clitico viene prodotto nella forma singolare: *i panini *l'ha salvato il cane*).
Per quanto riguarda i verbi eccezionali, come si è visto in fig. 2, le difficoltà sono effettivamente maggiori rispetto a quelle dei passivi e delle topicalizzazioni. Questo effetto è presente nel gruppo considerato globalmente, nonostante non si riscontri in Ida e Giulia, le due pazienti che presentano un quadro di afasia più grave, e che, oltre a non produrre correttamente alcuna topicalizzazione, non riescono a produrre correttamente neppure una frase passiva (v. fig. 3). La frase prodotta più spesso in modo corretto ha entrambi i nomi non marcati rispetto al numero (*a Vera manca l'ombrello*). In ogni caso, gli errori sono costituiti al 60% dall'omissione della marca del dativo (per es., **i ragazzi piace il pallone*), e al 78% dal verbo non accordato con il soggetto plurale (per es., *a Marco *manca i capelli*). Anche nel caso dei verbi eccezionali, come in quello dei passivi, dunque, gli errori possono essere interpretati come una tendenza a mantenere non marcate le corrispondenze tra le tre strutture, e dunque a interpretare come soggetto l'elemento più alto della gerarchia argomentale in prima posizione.

6 Conclusioni e prospettive

Questo lavoro presenta i primi risultati di una ricerca volta a verificare se le ipotesi elaborate dalla Teoria della Processabilità per l'acquisizione dell'italiano L2 sono applicabili alle condizioni di perdita afasica, nella prospettiva più generale di verificare, da un lato, se la TP può fornire il quadro teorico per contribuire alla messa a punto di strumenti per la valutazione e riabilitazione dei disturbi morfosintattici nei soggetti afasici e, dall'altro lato, se la perdita del linguaggio determinata dall'afasia può fornire dati a supporto alla TP. A questo scopo sono stati valutati sei soggetti afasici attraverso un protocollo volto ad elicitare tre strutture che nell'apprendimento risultano avanzate: i verbi passivi, la topicalizzazione dell'oggetto e alcuni verbi eccezionali.
Almeno ad una prima indagine, le ipotesi TP dello sviluppo implicazionale sembrano valere anche per i pazienti afasici: essi sono in grado di produrre frasi grammaticalmente corrette quando non è marcata la corrispondenza fra le tre strutture argomentale, funzionale e per costituenti. I problemi sorgono invece quando la corrispondenza è

marcata, poiché i pazienti tendono a regolarizzarla, e di conseguenza a produrre accordi errati, soprattutto quando la frase richiede marcato anche l'accordo per numero e genere. In questo senso, dunque, i risultati incoraggiano a proseguire l'indagine, cercando di raccogliere un corpus di dati più consistente, aumentando soprattutto il tipo di strutture e il numero di pazienti.

Nella prospettiva afasiologica, i task usati permettono non solo di evidenziare nei pazienti precise difficoltà che risultano invisibili ai test tradizionali, ma anche di individuare gerarchie implicazionali nel grado di difficoltà. Ciò indica che la TP può costituire una base teorica effettivamente utile per la costruzione di strumenti di valutazione dei disturbi grammaticali e per l'elaborazione di strategie mirate e graduate in base alla difficoltà per il trattamento riabilitativo dei soggetti che presentano difficoltà con più tipi di strutture. A questo scopo risulta opportuno proporre lo stesso protocollo a soggetti normali che abbiano l'italiano come L1 e che siano confrontabili agli afasici per età, scolarità e variabili socioculturali, così da avere dei valori con i quali confrontare le prestazioni dei pazienti.

Possiamo insomma concludere che questa prima esplorazione ha fornito dati sufficientemente incoraggianti per portare avanti questa interazione fra linguistica acquisizionale e afasiologia linguistica in una ricerca che si sta rivelando fruttuosa per entrambe le prospettive.

Riferimenti bibliografici

Bettoni, C. and B. Di Biase (in press) Processability Theory and its theoretical bases, in C. Bettoni and B. Di Biase (eds) *Processability Theory: Current issues in theory and application*, Amsterdam: Benjamins.

Bettoni, C., B. Di Biase and S. Ferraris (2008) Sviluppo sintattico e sviluppo morfologico: ipotesi di corrispondenze nella Processabilità dell'italiano L2, in G. Bernini, L. Spreafico and A. Valentini (eds) *Competenze lessicali e discorsive nell'acquisizione di lingue seconde*, Milano: Franco Angeli, pp. 355-82.

Bettoni, C., B. Di Biase and E. Nuzzo (2009) Postverbal subject in Italian L2 – a Processability Theory approach, in D. Keatinge and J.-U. Keßler (eds) *Research in Second Language Acquisition: Empirical Evidence Across Languages*, Newcastle upon Tyne: Cambridge Scholars, pp.153-73.

Bettoni, C. and G. Ginelli (in press) Acquiring constituent questions in Italian as a second language, in C. Bettoni and B. Di Biase (eds) *Processability Theory: Current issues in theory and application*, Amsterdam: Benjamins.

Bresnan, J. (2001) *Lexical-Functional Syntax*. Oxford: Blackwell.

Caplan, D. and N. Hildebrandt (1988) *Disorders of syntactic comprehension*. Cambridge (MA): MIT Press.

Caramazza, A. (1988) 'Some aspects of language processing revealed through the analysis of acquired aphasia: The lexical system'. *Annual Review of Neuroscience* 11, pp. 395-421.

– (1997) 'How many levels of processing are there in lexical access?'. *Cognitive Neuropsychology* 14, 1, pp. 177-208.

Caramazza, A. and A.E. Hillis (1990) 'Where do semantic errors come from?'. *Cortex* 26, pp. 95-122.

Caramazza, A. and G. Miceli (1991) 'Selective impairment of thematic role assignment in sentence processing'. *Brain and Language* 41, pp. 402-36.

Ciurli, P., P. Marangolo and A. Basso (1996) *Esame del linguaggio-II*. Firenze: Organizzazioni Speciali.

Di Biase, B. (2007) *A Processability approach to the acquisition of Italian as a second language: Theory and applications*. Tesi di dottorato, Canberra, Australian National University.

Di Biase, B. and S. Kawaguchi (2002) 'Exploring the typological plausibility of Processability Theory: Language development in Italian second language and Japanese second language'. *Second Language Research* 18, 3, pp. 272-300.

– (in press) Development across languages: English, Italian and Japanese, in C. Bettoni and B. Di Biase (eds) *Processability Theory: Current issues in theory and application*, Amsterdam: Benjamins.

Favilla, M.E. (2003) *Il linguaggio afasico nell'adulto. La linguistica aplicata alla valutazione e alla riabilitazione*. Tirrenia (Pisa): Edizioni del Cerro.

Favilla, M.E. and L. Ferroni (eds) (2009) *Disturbi del linguaggio e Neurolinguistica*. Perugia: Guerra Edizioni.

Favilla, M.E., L. Ferroni and M. Giannini (2009) I disturbi dell'elaborazione pragmatica, in M.E. Favilla and L. Ferroni (eds) *Disturbi del linguaggio e Neurolinguistica*, Perugia: Guerra Edizioni, pp. 121-43.

Ferroni, L. (2009) Basi teoriche nella riabilitazione dei disturbi sintattici, in M.E. Favilla and L. Ferroni (eds) *Disturbi del linguaggio e Neurolinguistica*, Perugia: Guerra Edizioni, pp. 99-120.

Friedmann, N. (2002) 'Question production in agrammatism: The tree pruning hypothesis'. *Brain and Language* 80, pp. 160-87.

– (1997) 'Tense and agreement in agrammatic production: pruning the syntactic tree'. *Brain and Language* 56, pp. 397-425.

Grodzinsky, Y. (1986) 'Language deficits and the theory of grammar'. *Brain and Language* 27, pp.135-59.

– (2000) Overarching agrammatism, in Y. Grodzinsky (ed.) *Language and brain. Representation and processing*, San Diego: Academic Press, pp. 73-86.

Hillis, A.E. and A. Caramazza (1991) 'Mechanisms for accessing lexical representations for output: evidence from a category-specific semantic deficit'. *Brain and Language* 40, pp. 106-44.

– (1995) 'Converging evidence for the interaction of semantic and sublexical phonological information in accessing lexical representations for spoken output'. *Cognitive Neuropsychology* 12, pp. 187-227.

Levelt, W.M. (1989) *Speaking. From Intention to Articulation*. Cambridge (MA): MIT Press.

Luzzatti, C. and R. De Bleser (1996) 'Morphological processes in Italian agrammatic speakers: Eight experiments in lexical morphology'. *Brain and Language* 54, pp. 26-74.

Miceli, G., B. Benvegnù and R. Capasso (1997) 'The independence of phonological and orthographic lexical forms: evidence from aphasia'. *Cognitive Neuropsychology* 14, pp. 35-69.

Miceli, G. and R. Capasso (1997) 'Semantic errors as neuropsychological evidence for the independence and the interaction of orthographic and phonological word form'. *Language and Cognitive Processes* 12, pp. 733-64.

Miceli, G., R. Capasso and A. Caramazza (1994) 'The interaction of lexical and sublexical processes in reading, writing and repetition'. *Neuropsychologia* 32, pp. 317-33.

Patterson, K.E. (1986) 'Lexical but non-semantic spelling?'. *Cognitive Neuropsychology* 3, pp. 341-67.

Pienemann, M. (1998) *Language processing and second language*

development: Processability Theory. Amsterdam: John Benjamins.
Pienemann, M., B. Di Biase and S. Kawaguchi (2005) Extending Processability Theory, in M. Pienemann (ed.) *Cross-linguistic aspects of Processability Theory*, Amsterdam: John Benjamins, pp. 199-252.
Pinker, S. (1984) *Language Learnability and Language Development.* Cambridge (MA): Harvard University Press.
Schwarz, M.F., R.B. Fink and M. Saffran (1995) 'The modular treatment of agrammatism'. *Neuropsychological Rehabilitation* 5, 1/2, pp. 93-127.

Il parlato schizofrenico
Una proposta di analisi conversazionale per seguire le piste topicali

Annamaria Cacchione
Università di Roma "Sapienza"
annamaria.cacchione@gmail.com

Abstract

This paper shows the results of a study focused on the topic management in semi-spontaneous spoken speech by schizophrenic patients. Three main issues have been investigated in order to trace the "topic path" of each conversation: the coherence and cohesion mechanism, the turn-taking pattern and the argumentation structure.
The study adopted a textual (i.e., not limited to the clause boundaries) definition of topic. Six therapeutic conversations involving four patients were selected among a corpus of psychiatric conversations collected by the CIRASS Department of the University of Naples "Federico II".
The findings show a great variety of phenomena, that can be traced back to common factors only in part. They also show a severe deficit in following the cooperation principle; as a consequence, the usual conversational patterns are often broken. In one case, the patient's speech appears to be absolutely normal (the speech is never disorganized) but shows different kinds of problems, mostly due to the reality perception.
It is suggested that a pragmatic deficit can account for many aspects of schizophrenic patients' language problems. A theory of mind related deficit can also account for more general impairments causing indirectly speech disorganization.

1 Introduzione

Questo contributo è il resoconto di una ricerca volta a studiare il trattamento del topic da parte di un piccolo gruppo di soggetti schizofrenici nell'ambito di colloqui terapeutici.
L'interesse per il parlato di soggetti con patologie psichiatriche, e schizofrenici in particolare, ha dato vita negli ultimi anni ad un ambito di ricerche internazionali cospicuo ed articolato (cfr. ad es. Fraser et al. 1986, Tènyi et al. 2002, Foundon et al. 2004, Covington et al. 2005), costantemente alla ricerca di definizioni e metodologie investigative capaci di rendere conto di fenomeni tanto complessi

quanto sfuggenti alle tradizionali categorie di analisi. Per quanto riguarda il panorama italiano, la ricerca si innesta in un più recente filone di studi dedicato al parlato patologico[1] e condotto con metodologie di analisi conversazionali (Bercelli, Leonardi e Viaro 1998), ricollegandosi direttamente ad altre ricerche realizzate negli ultimi due anni sullo stesso corpus di riferimento (Dovetto-Gemelli, 2008), l'unico che attualmente mette a disposizione, oltre all'audio ed alle trascrizioni, anche dati importanti sulla storia clinica dei soggetti.
L'attrattività e, insieme, la difficoltà di questo genere di studi deriva, da un lato, dalla difficoltà intrinseca dello studio del parlato spontaneo in generale e, dall'altra, dai problemi legati alla definizione di patologie mentali ancora oscure per molti aspetti, le cui manifestazioni linguistiche, dirette e indirette, sono oggetto di tentativi continui di analisi e sistematizzazione. In questo senso, anche in assenza di un quadro di riferimento completo, partire proprio dall'espressione linguistica patologica assume il valore di una strategia di carattere jakobsoniano, per cui l'analisi del singolo caso "deviante" può gettare luce sul funzionamento normale della struttura proprio a partire da ciò che non funziona come dovrebbe.

2 Il corpus

La ricerca è stata condotta su di un corpus di parlato raccolto dal CIRASS dell'università degli studi di Napoli "Federico II" e dalla Scuola sperimentale per la formazione alla psicoterapia e alla ricerca delle scienze umane applicate della ASL Napoli 1.
Il corpus, ancora in fieri, è attualmente composto da 17 ore di parlato prodotto nell'ambito di dieci colloqui terapeutici, in ognuno dei quali cui è presente un paziente e uno o più psichiatri.
I colloqui sono stati audio registrati e poi trascritti ortograficamente secondo i criteri adottati dai progetti API e CLIPS. Questi criteri, relativamente larghi (ad es. per le pause è stata adottata un'unica suddivisione in pause lunghe <lp> e brevi <sp>), favoriscono l'omogeneità di lettura e trattamento rispetto ad altri corpora e rendono inoltre più agevole l'adattamento a tipi diversi di analisi.

1 Risale al 1967 il lavoro, per tanti aspetti pionieristico, di Piro sul linguaggio schizofrenico. Gli anni '90 vedono il contributo importante di Pennisi (1994, 1997).

Il corpus, presentato da Dovetto e Gemelli nell'ambito del Convegno sulla Comunicazione parlata del febbraio 2006, è stato già oggetto di alcune analisi che hanno indagato aspetti diversi del parlato schizofrenico. E' inoltre in fase di allestimento uno spazio web nel quale saranno contenuti i dati e i risultati di tutte le ricerche condotte sul corpus[2] (in www.psicoterapia-scuola.it).
Per questa ricerca sono stati prese in esame 6 conversazioni, nei quali compaiono 4 pazienti (A, B, C e D). I pazienti B e C compaiono in 2 conversazioni ciascuno (B01, B03, C01 e C03). I pazienti A e D compaiono invece soltanto in una conversazione (A01 e D01). L'esposizione dei dati più significativi riguarderà tuttavia 4 conversazioni, una per ciascuno dei soggetti coinvolti.

3 Nozioni operative e metodologia di analisi

L'analisi è stata condotta secondo una prospettiva essenzialmente *corpus driven* (Chiari 2007), cioè senza elaborare preliminarmente ipotesi (da verificare poi attraverso l'analisi dei dati), ma determinando in via preliminare soltanto l'ambito di analisi, per poi pervenire, attraverso l'analisi stessa, all'(eventuale) identificazione di regolarità e dunque alla formulazione di ipotesi descrittive, da confrontare, infine, con i risultati raggiunti da altri tipi di analisi condotte in ambito linguistico e/o psicologico e psichiatrico.
Dal punto di vista operativo, si è trattato di ricercare nelle conversazioni tracce delle modalità adottate dai Follower (pazienti) per gestire il topic affidato loro dal Giver (terapeuta) oppure proposto autonomamente.
Con l'espressione *trattamento del topic* si comprende tutto ciò che rientra nei meccanismi di gestione del topic, inteso come principale argomento di riferimento del discorso (vedi oltre) nell'ambito, in questo caso, della dinamica conversazionale tipica del genere testuale "colloquio psicoterapeutico", che prevede, solitamente, la conduzione della conversazione da parte del Giver, ma non tanto stretta e predeterminata da annullare il carattere di spontaneità del parlato del

[2] Il corpus è attualmente disponibile presso il CIRASS dell'Università degli studi di Napoli "Federico II", la Scuola sperimentale per la formazione alla psicoterapia e alla ricerca delle scienze umane applicate della ASL Napoli 1

Follower. In questo senso, la gestione del topic da parte del Follower può essere ricondotta ai seguenti punti di snodo:

a. instaurazione del topic (se concesso da parte del terapeuta/Giver)
b. sviluppo del topic proposto dal Giver (che propone il topic sotto forma di domanda oppure, direttamente, di topic da "svolgere" – esempio: "parlami di…")
c. cambiamento e/o annullamento del topic in corso di svolgimento oppure (ritenuto) esaurito.

Il topic è tuttavia una nozione problematica e non univoca, che riunisce in sé aspetti testuali, sintattici, psicologici e cognitivi (Signorini 2003). Il topic come argomento, "ciò di cui si parla", si confonde infatti con le nozioni contigue di soggetto (soprattutto psicologico), tema (secondo la prospettiva strutturale), dato (in prospettiva informativa) e presupposto (secondo la posizione chomskiana), contribuendo a definire le altrettanto sfumate dicotomie di topic-comment, tema-rema, dato-nuovo e presupposto-focus.
Recentemente, in ambito italiano, è stata avanzata la proposta di una definizione trifattoriale del topic a sfondo essenzialmente pragmatico. Questa definizione, proposta da Cresti (1987) ed applicata nel corpus LABLITA, associa un criterio funzionale, uno distributivo e uno di tipo prosodico: il topic è un'unità informativa/locutiva che definisce il campo applicativo del comment (che è invece un'unità illocutiva) e ne crea le premesse semantiche; occupa una posizione tendenzialmente definita e stabile (precede il comment) ed è contraddistinto da un peculiare profilo intontivo[3].
Si tratta comunque di definizioni che pertengono essenzialmente all'ambito frasale, cioè del singolo enunciato, mentre mancano quasi del tutto nozioni operative capaci di orientare l'analisi del topic in ambito sovrafrasale, ed in particolare nell'ambito di testi ampi e complessi come quelli conversazionali.
Per ricostruire dunque i percorsi che il locutore ha compiuto nel suo lavoro di manipolazione del topic in relazione agli scopi sia suoi che dell'interlocutore, l'analisi è stata rivolta ad aspetti appartenenti a domini diversi ma interrelati come i seguenti:

[3] Esempio tratto da Signorini (2003): *impianto* /^{top} *che vuol dire* ?^{com} .

- strategie di coerenza, intesa come continuità topicale e analizzata soprattutto nelle forme di integrazione e coerenza relazionale (Bazzanella 2005, Givòn 1983);
- strategie di coerenza, intesa come salienza sintattica, testuale, situazionale e cognitiva (Bazzanella 2005);
- strategie di coesione come meccanismo elettivo di garanzia della coerenza, analizzato in riferimento all' articolazione delle catene coesive, soprattutto attraverso forme pronominali e all'uso di connettivi (Simone 1990, Bazzanella 2005);
- dinamiche conversazionali in relazione al tipo specifico di testo (colloquio terapeutico, caratterizzato dunque da asimmetricità anche se, nei casi in esame, non troppo evidente, vd. Orletti 2000, Leonardi-Viaro 1998);
- dinamiche conversazionali generali, per quanto riguarda soprattutto le capacità cooperativa del Follower (con riferimento esplicito ai principi di Grice 1975 ed alla cortesia[4] di Lakoff 1973).

E' possibile riassumere quanto sopra indicato in una griglia di osservazione che focalizza gli aspetti e gli item da indagare. Questa griglia riprende e declina in domande-chiave gli ambiti e le nozioni sopra elencate, compilando una sorta di check-list che è stata utilizzata, in senso largo e non prescrittivo, come guida all'esplorazione delle conversazioni.

A	*Il turno del Follower riprende effettivamente il topic proposto dal Giver?*
B	*Il turno del Follower riprende/sviluppa (eventualmente il topic del Giver) nella direzione richiesta e in modo adeguato?*
C	*Se il topic appare gestito in maniera non adeguata o regolare, è possibile rintracciare altre strategie (sintattiche, pragmatiche, cognitive ecc.) messa in atto dal Follower?*
D	*E' possibile raggruppare i topic "devianti" secondo qualche criterio?*
E	*Come vengono usati i connettivi?*

[4] Il principio di cortesia è qui inteso nella sua accezione più ampia di non conflitto con l'interlocutore, e cioè, per quanto riguarda la gestione dell'interazione, nella disponibilità a conversare ed in particolare a rispondere alle domande del Giver.

F	*Come vengono usati i pronomi, i deittici ed in genere gli altri coesivi?*
G	*Emergono altre peculiarità dalla conversazione del Follower, anche se non in diretta relazione con la gestione del topic?*

Tabella 1: check-list per l'analisi della gestione del topic

4 Analisi dei dati

4.1 Soggetto A: tra il mondo antecedente e il mondo conseguente

La conversazione esaminata (codificata come DGpsA01Nb) presenta un topic principale facilmente identificabile, instaurato chiaramente all'inizio del colloquio in riferimento al colloquio precedente.

(1) G#1: allora <lp> buongiorno <lp> senti ******* che <ehm> che cosa ci eravamo / #<F#2> avevi# scelto come argomento per oggi ?
F#2: #<G#1> <eeh># {<NOISE> <lp> parlare di quelle persone parland+ <eeh> passando dal mondo antecedente <sp> da / cioè dal mondo co+ <lp> antecedente al mondo conseguente diventano più umane

Per tutta la durata del colloquio (48 minuti) il passaggio "dal mondo antecedente a quello conseguente" sarà dunque il tema dominante.
L'andamento della conversazione è generalmente buono: il Follower risponde alle domande, anche se in modo un po' stringato. Non è mai scortese, e quando è oscuro non sembra farlo apposta. Non dà mai più informazioni del richiesto, talvolta meno del dovuto, spesso, anzi, le sue risposte appaiono appena sufficienti.
La coerenza appare complessivamente preservata. Non ci sono casi clamorosi di topic incoerenti (inseriti senza ragione o sviluppati in direzioni non coerenti), tuttavia in tre luoghi, nell'ambito di turni complessivamente ben organizzati, emergono dei rapporti logici particolari, definibili di "causalità inversa":

(2) G#51: sei riuscito a l+/ ad uscire<ee> di casa a fare una passeggiatina <sp> oppure no ?
F#52: <lp> <tongue click> il fatto è che poi piove {[whispering] se mi allontano

(3) G#65: e tu rientri subito <sp> appena hai gettato la spazzatura oppure<ee> ti trattieni un po'<oo> fuori ?
F#66: <sp>no, no non mi trattengo molto <sp> il tempo di tornare

A volte si ha l'impressione che a orientare il discorso, ad esempio nella scelta di un elemento lessicale, sia l'aspetto fonico più che il significato di un lessema precedente, come nel caso seguente (associazione fonica: insito→insinuazione):

(4) G#249: parlare guardandosi <sp> e così, il piacere si manifesta <lp> sapresti dire di che tipo di piacere si tratta ?
F#250: <lp><tongue click> è il piacere d+ / insomma insito <sp> in noi<ii> <sp> quasi un'insinuazione

Per quanto riguarda la coesione, il Follower mostra una gestione generalmente corretta dei coesivi: deissi personale, aggettivi dimostrativi, connettivi (*dunque*, *quindi*, *cioè*, *insomma*). Da rilevare l'uso insistito di *insomma* (56 occorrenze).
Si notano molte ripetizioni di parte (soprattutto finale) del turno precedente del Giver. Questa modalità diventa il pattern principale di inizio turno di risposta. All'ascolto assume talvolta cadenza ecolalica.

(5) G#189: <mh> <lp> e tu che cosa pensi della tua sofferenza ?
F#190: <vocal> <lp> {[whispering] che penso} <sp> penso e ripenso <sp> senza una / un<nn> <lp> senza capir+ / no senza<aa> <lp> {[whispering] <tongue click> penso e basta} <lp> <inspiration> che ne penso <breath> <lp> {[whispering] che ne penso} <sp> grave colpa certo <tongue click>

Il parlato di A evidenzia un atteggiamento cognitivo particolarmente interessante, e cioè la difficoltà a entrare in relazione con il mondo esterno nelle modalità ritenute normali: ad esempio, per A è molto difficile uscire fuori casa senza avere uno scopo preciso e/o sapere con esattezza chi si incontrerà. La dimensione dell'imprevisto sembra causargli problemi e ansia. In modo analogo, la descrizione che dà al Giver delle modalità in cui si rende conto degli stati d'animo altrui mostra aspetti singolari (si accorge del piacere degli altri dai capelli).

(6) G#69: no ? <lp> <clear throath> <lp> allora dicevi però <lp> uscire <sp> semplicemente per fare una passeggiata <clear throath> è una cosa che <sp> ti sembra più difficile da<aa> da realizzare è vero ? <lp> cerca di<ii> spiegarmi per / aiutami a capire un po' meglio, perché è così difficile fare una passeggiata ?
F#70: perché<ee> <lp> <tongue click> insomma si<ii> <lp> <inspiration> {[whispering] perché si incon+} <lp> <inspiration> insomma non si sa mai chi si incontra pe+/chi c'è per strada come dire<ee> <lp> non si prevede<ee> la persona

in cui si incontra <lp> cioè la persona<aa> che sta per strada come te <sp> n+ […] <clear throath> insomma il fatto<oo> <lp> che la gente insomma<aa> non è s+ / non / insomma non ci sta sempre uno fisso lì<ii> ce ne stanno altri <lp>
G#81: e questo <sp> fatto che <sp> ci sono altri <clear throath> perché costituisce un impedimento <sp> all'uscire ?
F#82: <lp> è proprio<oo> / questa è la / una verità<aa> <lp>
G#83: non ho capito
F#84: è così <lp> #<G#85> se uno non sa# la persona che lo aspetta come fa insomma<aa> a uscire ? <lp> dove va ? <sp> dove va ?

(7) G#253: […] tu ti accorgi quando <sp> una persona prova <clear throath> prova un grande piacere ?
F#254: <vocal> <mh>
G#255: da cosa te ne accorgi ?
F#256: <vocal> <sp> da cosa me ne accorgo <sp> da / insomma da <ehm> dagli occhi pure<ee>
G#257: <mh>
F#258: <sp> dal colore<ee> <sp> degli occhi <sp> cioè non è tetro, nero insomma, ma è chiaro, luminoso <lp>
G#259: quindi dagli occhi <sp> poi, da che cos'altro te ne accorgi ?
F#260: e<ee> insomma anche dai<ii> dai capelli<ii> <lp> pettinati insomma [whispering] <sp> in ordine

4.2 Soggetto B: il caos topicale

La conversazione (DGpsB01Nb) è, tra quelle esaminate, la più difficile da definire per quanto riguarda l'individuazione dei topic, e, più ancora, per l'individuazione di un topic principale, o, almeno, di un topic iniziale dal quale si dipartano topic secondari, come è plausibile che accada in un colloquio terapeutico (oltre che in ogni altro tipo di conversazione). L'impressione generale è quella di un vero e proprio caos conversazionale, e, quindi, anche topicale. Anche il Giver, per la maggior parte dei casi, rinuncia ad intervenire, poiché nel turno del Follower B non trova appigli ai quali riallacciarsi. Per questo motivo, la conversazione si presenta come una successione di monologhi di B, intervallati ogni tanto da brevi interventi del Giver. B, inoltre, riempie i turni di espressioni ricorrenti, stereotipate, private del loro significato: *hai capito/e' capit'?* (oltre 170 occorrenze), *giusto?*, *insomma*, *eventualmente*. Ricorrono spesso anche espressioni come *non lo so* e *non mi ricordo*, e *va a finire* (per lo più senza *che*) viene usato come puro riempitivo o formula di transizione.

(8) F#10: […] po+ f+ fu fu più importante il dischetto che il disco <inspiration> tant'è vero che va a finire <vocal> accorsero tutti quanti a <inspiration> a non cancellare no ? il<ll> dischetto <sp> no ?

Diversi i casi di code shifting, anch'essi casuali e non espressivi. Molte le pause, i tongue click, gli allungamenti vocalici e i segregati verbali (inspirazioni).
Le poche volte che B risponde riprendendo effettivamente il topic proposto è verboso o comunque usa circonlocuzioni strane:

(9) F#2: <inspiration> dopo quest'incidente tutto bene
G#3: <mh>
<sp>
<NOISE>
F#4: <tongue click> il lavoro altrettanto bene <sp> senza *** va mai #<G#5> bene<ee> <sp> tutto il lavoro# <sp> <tongue click> forse sono pure invidiato da<aa> <inspiration> da<aa> altre persone
G#5: #<F#4> <clear throath># <lp> #<F#6> < perché ?#
F#6: #<G#5> <inspiration> <tongue click>

Volendo dunque provare a rintracciare un topic principale, dopo un riferimento iniziale – un computer nuovo che B non sa usare – che sembra abbastanza chiaro si perde quasi subito ogni riferimento perché B passa subito ad altro. In una intera conversazione di circa 55.000 parole, il computer ricorre solo 20 volte.
Il brano che segue è esemplare dell'andamento della conversazione. Nella prima parte (primi 4 turni) sembra esserci una certa continuità topicale (si parla di computer e dischetti) anche se i nessi non sono chiari (*fare il disco nel dischetto*) e si nota l'intrusione di elementi non pertinenti e/o non spiegati (*il vanto di films*). Al turno successivo i nessi tra i topic sono ancora più labili fino a scomparire: il dischetto era talmente importante che B non ricorda cosa contenesse; B (che in precedenza ha detto di non saper usare il computer) ha inventato un linguaggio di computer del quale ha approfittato Bill Gates; da tutto questo è nata la ferrovia; poi in televisione è comparso Bassolino e ha rivelato tutto il piano. Quando il G. chiede cosa c'entri il dischetto con la ferrovia B. ritorna al computer. Si nota inoltre l'uso opaco dei pronomi (*loro*), il cambio ingiustificato della persona (*loro-c'aveva*), l'inserzione improvvisa di un discorso diretto rivolto ad un tu ignoto.

(10)[5] F#6: […] io veramente feci un fatto del genere tra parentesi questo non lo sa nessuno te lo dico a te 'e capi'
G#7: be' cosa facesti?
F#8: i dischetti questo in America no?
G#9: mh
F#10: e fu un vanto di films che fecero cioè in pratica non feci altro che fare il disco nel dischetto cioè fu più importante il dischetto che il disco tant'è vero che va a finire accorsero tutti quanti a non cancellare il dischetto no? molti di quei dischetti valse parecchio quel dischetto che tutti quanti c'andavano appresso giusto? e fu una caccia al dischetto giusto?
G#11: ma cosa conteneva questo dischetto?
F#12: non mi ricordo
G#13: mh
F#14: non mi ricordo adesso […] ho fatto un linguaggio che adesso chi è? Questo scemo che s'è fatto i soldi Gates no? […] quello è il problema per loro non era il problema no? per loro non era il problema perché tanto c'aveva lo stronzetto che teneva le idee 'e capi' come ne è nata la ferrovia 'e capi' tant'è vero che *** giust'? compagno di laboratorio mio, "simm' asciut' nui ca' ferrovia 'amm capit' tutt' cos' chell' che vuliv' fa' tu" hai capito ch' co' uno stronzetto questo in anticipo rispetto a poi dopo Bassolino che poi s'è visto per televisione ha detto tutto il piano da fare diciamo ora manc' solamente per la verità di tutto il taglio del nastro e poi il via insomma ai lavori insomma questo quando sarà se sarò vivo vedrò qualche cosa non lo so 'e capi' capit' mo' mancava la ferrovia […]
G#19: che c'entra la ferrovia con il dischetto?
F#20: no dico perché nel momento in cui noi programmavamo e serviva un linguaggio di computers […]

Talvolta sembra di rintracciare un collegamento tra i topic introdotti nella somiglianza fonica dei lessemi e/o nella vicinanza di significato delle accezioni dello stesso tipo di lessemi, come nella catena: *rappresentare questa situazione → chi ti rappresenta? → rappresenta il parassita* (forse evocato dalla parola microbo detta all'inizio) → *rappresenta schematicamente le persone → ti schematizza*

(11) F#34: […] tutto questo accadde nel momento in cui zio *** e tanti di noi rappresentammo *questo situazione alle elementari io ne parlai co' zio *** e mi

[5] Per questo esempio e per il successivo la trascrizione è stata normalizzata perché quella originaria risulta poco leggibile per la presenza pervasiva di code shifting e fenomeni paralinguistici (es.: non so se mi spiego <lp> so caspiti s+ / <dd>de+ dei suoi <lp> {<dialect> 'e capi' ? </dialect>} <sp> {<whispering> io non c'entro</whispering>} <lp> poi {<dialect> pe' mme </dialect>} <sp> sta bene così).

disse quello a te chi ti rappresenta ? zio *** mi disse che quello rappresentava il parassita tu a quello non lo devi proprio vedere anzi quando lo vedi
G#35: mh
F# 36: giusto ? quando lo vedrai nun l' 'e 'a vere' proprio 'e facc' l' 'e 'a schifa' cioè io a quel momento che facevo ? 'e capi' ? c'aveva fa' ? l'aveva sputa' 'n facc' e c'aveva ra' 'nu cavec' 'n cul' 'nsomm e dicev' vavattenn' 'e capi'?
G#37: ma chi era questo parassita di cui parlava zio *** ?
F#38: no questo non l'ho capito cioè lui rappresentava schematicamente le persone cioè come dire tu sei questo ? no ? e ti schematizzava per questo

4.3 Soggetto C: gli organi semoventi, ovvero come gestire in modo corretto topic assurdi

Le conversazione con il soggetto C si distinguono dalle altre perché si presenta formalmente del tutto regolare.
Le sequenze di turni sono organizzate in modo corretto, e la conversazione che ne deriva è molto scorrevole e piacevole. Si ha l'impressione che anche il G sia coinvolto positivamente nello scambio e si appassioni a quanto C dice. Anche l'ascolto conferma questa scorrevolezza: il tono è vario, la fluenza buona, la prosodia coerente con la struttura sintattica e semantica degli enunciati. Le pause sono per lo più brevi e corrispondono all'articolazione della pianificazione tipica del parlato.
I principi di cooperazione sono dunque tutti rispettati, così come quelli di cortesia sono pienamente rispettati.
Il linguaggio è lessicalmente ricco, a volte anche ricercato, sintatticamente articolato e non banale. Notevole inoltre l'uso di molto termini appartenenti a linguaggi settoriali di tipo tecnico-scientifico (anatomia, fisiologia, elettronica). C inoltre usa esclusivamente l'italiano e non ci sono casi di code shifting, tranne in qualche raro caso in cui il dialetto è usato con funzione espressiva e limitata ad esclamazioni tipiche (*mannaggia*). La coesione è assicurata da un uso corretto e chiaro dei referenti e delle catene coesive. Negli esempi che seguono, C si rende conto del fatto che un coesivo potrebbe avere una referenza ambigua e provvede ad esplicitarla, spontaneamente nel primo caso, su richiesta del G. nel secondo:

(12) G#343: non ti accade nulla che tu non voglia

F#344: ah sì se lui l'entità che è dentro di me decide di scherzare non me le fa usare nemmeno

(13) F#564: no, perché stranamente mi parla quando lei non c'è sta face+ / cercano di fare in modo che io non parli più davanti a loro non mi interrogano più davan+ / da quando ci sono loro
G#565: non ti interrogano più chi ?
F#566: sia *** che questa voce

Anche l'uso dei connettivi è regolare. C appare inoltre in grado di gestire senza problemi meccanismi di spostamento della referenza verso il contesto (salienza situazionale) senza interferire con il co-testo conversazionale.

(14) (Durante la pausa lunga il Follower si accende una sigaretta con l'accendino che evidentemente gli porge il dottore, così lo ringrazia)
G#421: prego
F#422: è un'ottima cosa il registratore dottore, ma sta registrando ancora !
G#423: sì
F#424: quando la vogliono fare questa legge dottore quando pensate che si potrebbe fare? […] perchè questa voce mi ha detto che io sono sempre collegato con *** anche se parlo solo c'è sempre un collegamento come se come un circuito che parte dalla presa e va nel registratore

La coerenza come integrazione, capacità di organizzare i topic e i sub-topic in sequenze causali e temporali corrette, è più volte manifestata in modo evidente, come nella ricetta del ragù.

(15) F#638: allora si mette cipolla olio poi si prende la carne macinata si fa rosolare e poi si aggiunge la salsa la salsa liquida la passata e si fa cuocere un paio d'ore a fuoco lento
G#639: e viene molto buono?
F#639: ottimo, straordinario perchè piglia un sapore bellissimo quando cuoce così

Dov'è allora il problema?
Il problema è che molte argomentazioni, pur essendo ben costruite, sono fondate su premesse assurde, perché corrispondenti a fatti impossibili che vengono però vissuti e percepiti come reali o almeno plausibili. Tutta la parte centrale della conversazione, relativa alle funzioni degli "organi cerebro-semoventi" di cui C si dichiara dotato, è strutturata in questo modo.

(16) F#24: se ci sono (Nota: gli organi cerebro-semoventi) è perchè<ee> servono a uno scopo <lp> se una cosa esiste è perchè occorre <sp> deve avere uno scopo, oppure no?

(17) F#348: <tongue click> gli zigomi l'ho detto <sp> <mh> <sp> <tongue click> e il radar al centro del cervello <sp> il radar al centro del cervello eccolo qua <lp> questo qua <mh> <sp> è il cervello <lp> e il radar parte dal centro <NOISE> <sp> è un raggio che io mando sopra <sp> e poi <sp> faccio così lo faccio girare <sp> in senso orario e antiorario
G#349: <mh>
F#350: così <sp> e forma un cono, anche questo
G#351: <mh>
F#352: solo che questo raggio io lo posso dirigere <lp> lo posso dirigere in questo modo <lp><NOISE> […]
G#355: e anche per questo <sp> non<nn> hai nessuna idea degli #<F#356> effetti#
F#356: #<G#355> no perché# <sp> se fosse un radar io dovrei localizzare le nuvole gli aeroplani tutto ciò che vola

E' particolarmente interessante notare come, dall'interno di una delle argomentazioni più importanti nell'economia dell'intero colloquio – quella in cui viene affrontato il topic delle voci che C sente - C sembri addirittura essere in grado di esercitare un'azione di critica e di poter vedere la situazione da prospettive diverse dalla sua. Pur non mettendo mai seriamente in discussione la realtà della situazione descritta, C è però capace di parlarne con sottile umorismo – il che implica sempre un certo distanziamento dall'oggetto. Inoltre, C riconosce alle voci un ruolo importante nella sua crescita personale: costringendolo ad allenare il cervello per affrontare la nuova situazione, lo hanno reso più sensibile, aumentando le sue capacità di comprendere le intenzioni ed i pensieri degli altri – la sua capacità di *mindreading*.

(18) G#411: e quante ne hai contate ?
F#412: <eeh> <sp> <eeh> siamo siamo a<aa> <sp> ventiquattro venticinque voci <sp> fino ad ora
<note> il '<eeh>' a inizio turno è piuttosto un'interiezione per intendere una quantità molto grossa </note>
G#413: <inspiration> quindi ventiquattro venticinque voci che<ee> <sp> rappresentano
F#414: sempre la stessa entità sempre <unclear>
G#415: è sempre la stessa #<F#416> persona#

F#416: #<G#415> è sempre la stessa# persona <sp> che mi piglia in giro <sp> io l'assecondo però non è che<ee> <sp> dico no non scherzo <sp> io<oo> gioco anche così <sp> passo proprio il tempo così a giocare con lui
G#417: ed è una<aa> <NOISE> un passatempo piacevole o #<F#418> spiacevole ?#
F#418: #<G#417> incredibile perché# mi sta<aa> <sp> mi allena il cervello in effetti <sp> io prima *era di un'ingenuità incredibile dottore <sp> non capivo tante cose tante sfaccettature della memoria {<NOISE> umana </NOISE>}<sp>
G#419: <mh>
F#420: invece le<ee> le intendo adesso <sp> quando una persona mi parla o<oo> <sp> o fa qualcosa<sp> io m+ molte volte<ee> intuisco <sp> le sue intenzioni <sp> e dove vuole arrivare con le sue azioni <lp> <NOISE> è meglio <laugh>

4.4 Soggetto D: strane favole

Le favole alle quali si fa riferimento nel titolo sono quelle che il G chiede al soggetto D di raccontare nel tentativo di farlo parlare in modo continuativo per qualche minuto. Il soggetto non appare infatti in grado di sostenere una conversazione normalmente: i suoi contributi conversazionali sono per lo più insufficienti il Giver stenta a trovare topic che facciano procedere il colloquio.
C procede con lentezza e fatica, ripetendo spesso in maniera meccanica singoli lessemi o espressioni usate come stereotipi (*dipende*, *tutt' e' maner'*) oppure pezzi di turni precedenti – in genere le parti finali, soprattutto delle domande – con effetti ecolalici:

(19) G#263: sei contento ?
F#264: <breath> *oento

Il suo parlato è monotono e stentato, punteggiato di allungamenti vocalici, specie in posizione finale, di tongue click e di altri segregati verbali (inspirazioni). Gli enunciati hanno tutti la stessa intonazione, con picchi di volume seguiti da sillabe pronunciate a bassa voce e in modo poco accurato, quasi trascinando l'articolazione. C parla inoltre quasi esclusivamente in dialetto.

A volte si ha l'impressione che a orientare il discorso, ad esempio nella scelta di un lessema, sia l'aspetto fonico più che il significato di un lessema precedente, come nel caso seguente (associazione fonica: cappuccetto→cappuccino):

(20) G#199: <mh> <sp> <tongue click> senti vo+ vorrei chiederti una cosa tu sai<ii> <sp> #<F200> la favola di cappuccetto rosso# ?
F#200: #<G#199> <eh!> <breath> <tongue click># <sp> di cappuccetto rosso<oo>? <sp> *sso
#<G#201> <tongue click># <inspiration>
G#201: #<F#200> la conosci ?#
F#202: <eeh> <tongue click> cioè cappuccetto rosso va<aa> al bar e {<dialect> s'accatt' 'nu </dialect>} cappuccino

In diversi passaggi i nessi logici non sono quelli che ci si aspetta, come nell'esempio seguente, in cui C dice che quando era fidanzato aveva 22, 24 o 27 anni, la sua fidanzata aveva la stessa età ed era più grande di lui:

(21) G#93: quanti anni avevi quand'eri fidanzato ?
F#94: <tongue click> <tongue click> {<dialect> tenevo<oo> sev' giovan' <sp> poc' vintiruje vintiquatt' vintisett' </dialect>} <sp>
G#95: e la tua fidanzata quanti anni aveva ?
F#96: <tongue click> <sp> {<dialect> stess' età è cchiù gross' 'e me<ee></dialect>}

Il passaggio seguente esemplifica alcune principali caratteristiche del discorso di C: difficoltà marcata a fornire spiegazioni, tentativo di risolvere il compito conversazionale/cognitivo attraverso una risposta stereotipata (*sacc' e cos'*) oppure deviando verso aspetti secondari del topic (al rigo 6, invece di spiegare di cosa parla il programma, C precisa che la trasmissione avviene un altro giorno). Esemplare anche la costruzione del turno di risposta 12: a una prima parte di ripresa del turno precedente del G., come forma di comprensione ed accettazione della richiesta, segue una seconda parte, tautologica, che nuovamente evita la spiegazione dovuta. Significativo, infine, l'uso del coesivo (*un altro nome*), senza che sia stato in precedenza instaurato alcune referente parallelo.

(22)[6] G#119: qual è l'ultimo programma che hai visto ?
F# 120: pe' television' Scherzi A Parte
G#121: Scherzi A Parte ?
F#122: eh!
G#123: e di cosa parla questo programma ?
F#124: sacc' 'e cos' (un sacco di cose)

[6] Anche in questo caso e nel successivo gli esempi sono stati normalizzati per agevolarne la lettura, anche perché in alcuni turno è stato necessario aggiungere la traduzione dal dialetto).

G#125: me lo puoi spiegare ? perché io non l'ho mai visto
F#126: chill' mo' fa 'n'atu juorn' 'a *trammission' fa' n'atu journ' fa mo' Scherzi A Parte nun fa mo' fa n'atu juorn' (quello ora fa un altro giorno la trasmissione fa un altro giorno fa ora Scherzi a parte non fa ora fa un altro giorno)
G#127: e chi è il presentatore ?
F#128: nun m'arricord' 'n'atu nomm' (non mi ricordo un altro nome)
G#129: è un altro nome ?
F#130: 'n altro nome eh!
G#131: e cosa ti è piaciuto di questo programma ?
F#132: m'è piaciut' eeh quand' ha fatt' pe' television'

Nel tentativo di trovare argomenti che permettano a C di sviluppare la conversazione, il G gli chiede di raccontare le favole che conosce. Vengono fatti diversi tentativi, che però sortiscono tutti lo stesso effetto: C non è in grado di raccontare alcuna storia perché non è in grado di esplicitare passaggi complessi. La deviazione dal topic proposto, anche soltanto come strategie di evitamento, di solito avviene non ad inizio turno ma all'interno di esso. C appare cioè in grado di collegarsi in modo coerente al topic proposto attraverso l'applicazione di semplici meccanismi di ripetizione di parti salienti del turno precedente, tuttavia non è in grado di proseguire nel compito assegnatogli, in questo caso perché sembrano mancargli dei frame stabili e articolati di favola (anche il repertorio delle favole è incerto, tant'è vero che include anche Dracula). Quella che tenta infatti di narrare è solo un abbozzo di favola: a volte è in grado di proporre uno sfondo (X nel bosco che fa Y) ma non sa proseguire verso un evento principale e una risoluzione. Allora tenta di risolvere il problema ricorrendo

- alla solita formula inclusiva (*favole di tutti i tipi - dipende*),
- a deviazioni dal topic principale (*quando ero piccolo andavo a scuola*),
- a invenzioni narrative accostando pezzi di favole diverse a caso insieme a "brandelli" di possibili eventi,
- all'interruzione del compito come se fosse terminato.

(23) G#155: senti quand'eri più piccolo che favole conoscevi ?
F#156: quand' sev' pezzerill' ? (quando ero piccolo?)
G#157: mh
F#158: eeh quand' sev' pezzerillo jev' a scol' (quando ero piccolo andavo a scuola)
G#159: mh [..] e che favole conoscevi ?
F#162: 'e favole ?

G#163: mh
F#164: 'e favole 'e tutt' maner' (favole di tutti i tipi)
G#165: eh me ne dici una?
F#166: una? 'e favole 'e favol' 'e dipend' favol che ch' pensav' quand' jev' a scol' 'sti ffavol' lloc' eeh (una? Le favole le favole le dipende favole che che pensavo quando andavo a scuola queste favole qua eeh)
G#167: ci puoi raccontare una favola ?
F#168: una favola ? e 'na favola 'i Dracula ? (una favola di Dracula?)
G#169: e raccontacela
F#170: cioè Dracula / Dracula nel bosco ch coglie 'i fragole […] agg' saput' ricer' facil' (cioè Dracula/dracula nel bosco che coglie le fragole… l'ho saputa dire facilmente)[…]
G#179: e come va a finire questa questa favola ? ce la racconti ?
F#180: favol' ca favol' ca eeh favol' che Dracul' nel bosc' ver' nu can' ver' nu can' (favola che favola che eeh favola che Dracula nel bosco vede un cane vede un cane)

5 Osservazioni conclusive

Le analisi presentate permettono di trarre alcune conclusioni sugli aspetti peculiari della gestione del topic nel parlato patologico di soggetti schizofrenici. Inoltre, come è logico aspettarsi, l'osservazione del trattamento del topic permette di ricostruire l'andamento generale della conversazione, fornendo dati anche su una serie di aspetti correlati come la gestione dei turni, i meccanismi coesivi, le strategie di coerenza e la gestione dei sistemi locali dei significati.
Pur nella ristrettezza del campione, le conversazioni analizzate hanno fornito una mole notevole di dati e hanno mostrato, soprattutto, una grande diversità di fenomeni che appare arduo ricondurre ad un unico tipo di correlato linguistico collegato alla patologia di base – la schizofrenia. Ciononostante, dei tratti in comune paiono comunque esserci, e, nella varietà degli "stili conversazionali" dei singoli, suggeriscono piuttosto una diversa manifestazione di un nucleo comune di difficoltà.
Dal punto di vista della gestione globale della conversazione, tutti i soggetti analizzati appaiono in grado di seguire l'alternanza dei turni, anche se con un grado di partecipazione molto diverso. Anche il principio di cortesia è sempre rispettato.
Il principio di cooperazione, al contrario, è spesso disatteso: i contributi di alcuni soggetti in particolare (B e D) non rispettano praticamente mai le massime di quantità, relazione e modo, mentre

non si ha motivo di dubitare della loro sincerità (ma ovviamente non è possibile esserne certi) e dunque si può presumere che rispettino la massima della qualità.
Per quanto riguarda i meccanismi di coesione, spesso i referenti risultano opachi e i pronomi vengono usati in riferimento a referenti prima mai instaurati. L'uso dei connettivi è deviante rispetto al loro significato. La ripetizione può venire usata come strategia coesiva, ma più spesso il soggetto ripete formule stereotipate oppure ripete in modo ecolalico l'ultimo pezzo della parola del turno precedente oppure anche del proprio contributo (si noti soprattutto il soggetto C).
I nessi logici possono essere usati in maniera contraria al loro uso regolare, come fa ad esempio il soggetto A in vari casi di "causalità inversa". In questo quadro generale, i topic vengono talvolta associati in base alla somiglianza fonica, oppure secondo associazioni semantiche particolari, secondarie rispetto al significato centrale oppure assolutamente personali. Nei casi di maggiore deviazione da una gestione regolare e plausibile (come nel caso di B), i topic compongono "cumuli" secondo criteri che risultano inaccessibili sia all'interlocutore che all'analista.
L'eterogeneità dei fenomeni riscontrati, osservata anche in corpus di riferimento molto più ampi (Meilijson et al. 2004, Kuperberg and Caplan 2003) richiede probabilmente categorie non discrete, più "fini", di analisi, capaci di rendere conto di una complessità tale da far sembrare vano ogni tentativo di sistematizzazione. In questo senso, è opportuno però anche aggiungere che studi di tipo linguistico sul parlato dei soggetti schizofrenici, almeno in Italia, sono ancora molto pochi, e che il loro sviluppo potrebbe dare contributi preziosi all'individuazione di tratti comuni riconducibili a loro volta a elementi comuni dell'*impairment* cognitivo.
Di particolare interesse appare la contrapposizione tra la capacità di strutturare argomentazioni perfette, mostrata da alcuni soggetti come C – argomentazioni basate però su premesse impossibili - e l'incapacità quasi totale di spiegare spiegare/argomentare mostrata da altri. A tale proposito, sarebbe utile avere informazioni sulle capacità pregresse dei soggetti in esame, per verificare se hanno perso vs. preservato capacità che in precedenza avevano. In questo senso, sarebbe anche interessante verificare se esiste una correlazione con il grado di scolarizzazione dei soggetti esaminati, come sembra

suggerire, ad esempio, il lessico ricco e specialistico dell'unico soggetto in grado di conversare normalmente (C).
Infine, l'aspetto più propriamente cognitivo del problema.
È possibile che la "sconnessione topicale" sia dovuta a difficoltà di *mindreading*, come alcune ricerche sembrano suggerire (Brüne and Bodenstein 2004, McCabe et al. 2005)? Che cioè i soggetti non siano in grado di prevedere che, ad esempio, non esplicitando una referenza, che per loro è chiara, l'interlocutore non riuscirà a recuperarla?
Questo tipo di difficoltà, collegata, secondo alcuni, al mancato o imperfetto sviluppo della "teoria della mente" nei soggetti schizofrenici (vd. Brüne 2005 per una rassegna della letteratura in materia), potrebbe manifestarsi linguisticamente nell'incapacità di gestire la coesione tenendo presente la prospettiva dell'interlocutore, risolvendosi quindi in una sorta di linguaggio privato inaccessibile all'esterno.
Questa ipotesi, che potrebbe aiutare a ricondurre molti degli aspetti osservati ad un nucleo unitario di problemi, ha però bisogno di verifiche più ampie, possibilmente longitudinali e corredate da informazioni di supporto importanti, come quelle sulla condizione pregressa dei soggetti analizzati.

Riferimenti bibliografici

Bazzanella, C. (2005) *Linguistica e pragmatica del linguaggio*. Roma-Bari: Laterza.

Bercelli, F., P. Leonardi and M. Viaro (1998) *Cornici terapeutiche. Applicazioni cliniche di analisi dell'interazione verbale*. Milano: Raffaello Cortina Editore.

Brüne, M. (2005) '"Theory of Mind" in Schizophrenia: A Review of the Literature'. *Schizophrenia Bulletin* 31, 1, pp. 21-42.

Brüne, M. and L. Bodenstein (2005) 'Proverb comprehension reconsidered – 'theory of mind' and the pragmatic use of language in schizophrenia'. *Schizoprenia Research* 75, pp. 233-39.

Chiari, I. (2007) *Introduzione alla linguistica computazionale*. Roma-Bari: Laterza.

Covington, M. A., C. He, C. Brown, L. Naçi, J.T. McClain and S.B. Fjordbak (2005) 'Schizophrenia and the structure of language: The linguist's view'. *Schizophrenia Research* 77, pp. 85-98.

Cresti, E. (1987) L'articolazione dell'informazione nel parlato, in A.A.V.V. (eds) *Gli italiani parlati*, Firenze: Accademia della Crusca, pp. 27-90.

Dovetto, F. M. and M. Gemelli (2008) *Il parlato di soggetti schizofrenici*, in M. Pettorino, A. Giannini, M. Vallone and R. Savy (eds) *La comunicazione parlata, Atti del Congresso Internazionale* (Napoli 23-25 febbraio 2006) – Tomo II, eBook, Napoli: Liguori, pp. 1081-92.

Foundon, N., A. Reboul, P. Sabatier and M.C. Noël-Jorand (2004) The production of discourse in four subjects affected with psychosis. French Cognitive Programme 2000-03, Avail. on-line http://hal.archivesouvertes.fr/docs/00/02/90/96/PDF/Fouresab.pdf (1/ 9/09).

French Cognitive Programme 2000-03, avail. on-line http://hal.archivesouvertes.fr/docs/00/02/90/96/PDF/Fouresab.pdf (1/ 9/09).

Fraser, W.I., K.M. King, P. Thomas, P. and R.E. Kendell (1986) 'The Diagnosis of Schizophrenia by Language Analysis'. *The British Journal of Psychiatry* 148, pp. 275-78.

Givón, T. (1983) *Topic Continuity in Discourse. A Quantitative Cross-Linguistic Study*. Amsterdam-Philadelphia: John Benjamins.

Grice, P.H. (1975) *Logic and Conversation*, in P. Cole P. and J.L. Morgan (eds) *Syntax and Semantics*, Vol. 3, Speech Acts, New York: Academic Press.

Kuperberg, G. R. and D. Caplan (2003) *Language Dysfunction in Schizophrenia*, in R.B. Schiffer, S.M. Rao and B.S. Fogel (eds) *Neuropsychiatry*, Philadelphia: Lippincott Williams and Wilkins, pp. 444-66.

Lakoff, R. (1973) *The logic of politeness; or, minding your P's and Q's*, in C. Corum, T.C. Smith-Stark and A. Weiser (eds) *Papers from the Ninth Regional Meeting of the Chicago Linguistic Society*, Chicago: Chicago Linguistic Society, pp. 292-305.

McCabe, R., I. Leudar and P.G.T. Healey (2005) *What Do You Think I Think? Theory of Mind and Schizophrenia*, in B.G. Bara, L. Barsalou and M. Bucciarelli, M. (eds) *CogSci 2005 Proceedings*, Mahwah, New Jersey: Lawrence Erlbaum Associates, pp. 1443-8.

Meilijson, S.R., A. Kasher and A. Elizur (2004) 'Language Performance in Chronic Schizophrenia: a Pragmatic Approach'. *Journal of Speech, Language, and Hearing Research* 47, pp. 695-713.

Orletti, F. (2000) *La conversazione diseguale. Potere e interazione*. Roma: Carocci.

Pennisi, A. (1994) *Le lingue mutole. Le patologie del linguaggio fra teoria e storia*. Roma: La Nuova Italia Scientifica.

– (1998) La macchina schizofasica. Psicopatologia e filosofia del linguaggio, in F. Lo Piparo F. and D. Gambarara (eds) *Linguaggio e cognizione*, Roma: Bulzoni, pp.349-64.

Signorini, S. (2003) *Il* Topic*: criteri di identificazione e correlati morfosintattici in un corpus di italiano parlato*, in F. Albano Leoni, F. Cutugno, M. Pettorino and R. Savy (eds) *Atti del convegno nazionale "Il parlato italiano"* (Napoli, 13-15 febbraio 2003), CD-ROM, Napoli: M. D'Auria Editore.

Simone, R. (1990) *Fondamenti di linguistica*. Roma-Bari: Laterza.

Tényi, T., R. Herold, I.M. Szili and M. Trixler (2002) 'Schizophrenics Show a Failure in the Decoding of Violations of Conversational Implicatures'. *Psychopathology* 35, pp. 25-7.

La produzione verbale dei soggetti con Sindrome di Down: aspetti fonologici e morfologici

Patrizia Sorianello
Università degli Studi di Bari
p.sorianello@lettere.uniba.it

Abstract

The present study deals with the linguistic production of four children with Down Syndrome (aged 6-10 years). Each subject was audiotaped during verbal elicitation tasks. Both phonological and morphological abilities of the children are described paying particular attention to consonantal system and to free-standing morphemes. Our study differs from previous Italian research on Down Syndrome: for the first time the phonological description is based on a spectrographic analysis. The acoustic results reveal high percentage of substitution, assimilation processes and cluster reduction. A relevant delay in phonological development, with more errors in the production of later-developing sounds, is observed. Articulatory imprecision can be directly attributed to the oral and facial anomalies; in addition, general hypotonia and specific oral motor problems affect the coordination and the timing of articulatory gestures. The findings of the morphological analysis confirm that children with Down Syndrome omit many grammatical morphemes demonstrating particular difficulty with the production of the functional categories such as: articles, clitics, prepositions and auxiliaries. The omission is probably a sort of economy strategy: functional words are non semantic elements with scarce phonetic and prosodic weight.

1 Premessa

La Sindrome di Down è una delle sindromi poliformative congenite più frequenti, circa un caso su 800 nati vivi. La sua origine è genetica, giacché è la conseguenza di una alterazione del cromosoma 21. Nei soggetti Down, sono presenti tre cromosomi 21 anziché due, per questo la Sindrome è anche detta Trisomia 21 (cfr. Lubec, 2002).

La Sindrome di Down manifesta una serie di tratti peculiari; a livello fisico alcune caratteristiche la rendono immediatamente riconoscibile come: l'obliquità della rima palpebrale, l'occipite piatto, la microcefalia, ed ancora ridotta statura ed eccesso ponderale. In questa sede, ci preme sottolineare, tra i tanti, quegli aspetti che, in misura diretta o indiretta, incidono sulla fonazione. In primo luogo, andranno menzionate le anomalie che si rilevano a carico dell'apparato fono-articolatorio: cavità orale piccola, dismorfismo dei denti, lingua larga e muscolosa, palato ogivale. Ipoacusia e patologie respiratorie sono inoltre frequenti disturbi secondari che compromettono, almeno in parte, la comunicazione verbale. A tutto ciò va aggiunto, quale fattore generale, la presenza di una diffusa iperlassità articolare e di una ipotonia muscolare.

Il dato clinico più rilevante è tuttavia costituito da un significativo ritardo mentale, il quale diventa sempre più evidente con l'avanzare dell'età cronologica. Il ritardo cognitivo non si deve a lesioni o ad alterazioni neuroanatomiche, quanto piuttosto ad un'inadeguata maturazione delle strutture cerebrali. Diversi studi hanno dimostrato, anche mediante l'impiego di moderne tecniche di neuroimmagini, che il volume e la densità del cervello dei trisomici è ridotto, così come lo sono le connessioni sinaptiche e la mielizzazione di diverse aree cerebrali, lo sviluppo del sistema nervoso centrale appare invece rallentato (cfr. Jernigan and Bellugi, 1990; Wang *et al*, 1992).

Inevitabili sono le ripercussioni sul piano linguistico. La letteratura sull'argomento è molto vasta, sebbene eterogenea per impostazione metodologica. Ad oggi, i processi più indagati risultano quelli legati alla sfera morfosintattica e lessicale; di gran lunga inferiore è il numero delle ricerche condotte in ambito fonetico e fonologico. I soggetti Down manifestano un significativo *deficit* linguistico, la cui entità è di gran lunga superiore alla loro età mentale, elevata è tuttavia la variabilità interindividuale. Il linguaggio dei Down presenta diversi comportamenti peculiari, talora asimmetrici. L'aspetto più evidente è che non tutto il sistema risulta compromesso nella stessa misura; lo sviluppo linguistico, più che un percorso globale, sembra mostrare un grado maturazionale diversificato. Fin dai primi anni di vita, comprensione e produzione verbale sembrano seguire schemi evolutivi differenti; il versante produttivo appare inadeguato e meno sviluppato rispetto alla capacità di comprensione, sebbene questa divergenza si attenui nel tempo. Fra le cause ventilate, un'insufficiente

programmazione motoria o anche una atipica lateralizzazione cerebrale con «dissociazione tra percezione linguistica, processata dall'emisfero destro, e produzione linguistica, controllata dall'emisfero sinistro» (Nicolai, 2003: 199).
Il componente linguistico maggiormente inficiato è quello morfosintattico, seguito da quello fonologico: la produzione verbale delle persone Down mostra natura telegrafica e ripetitiva, poiché contrassegnata da elevata alterazione morfologica. Per contro, il livello lessicale ha, almeno nei primi anni di vita, uno sviluppo approssimabile a quello dei soggetti normodotati. La divergenza tra comportamento morfosintattico e lessicale ha suscitato l'interesse degli studiosi: sviluppo modulare dei componenti linguistici (cfr. Fowler, 1990) e sviluppo asincrono indotto da svantaggio selettivo (cfr. Vicari, Caselli and Tonucci, 2000; Vicari *et al*, 2002) sono le ipotesi interpretative più frequentemente perseguite.

2 Il disegno sperimentale

L'obiettivo primario di questa ricerca è quello di descrivere la competenza linguistica di giovani soggetti affetti da Sindrome di Down. L'indagine ha riguardato solo il versante produttivo; i livelli esplorati sono due: fonologico e morfologico. La scelta di questi componenti linguistici non è casuale; complessivamente, l'ambito fonologico ha ricevuto una minore attenzione da parte della comunità scientifica; questo settore di indagine appare ancora più lacunoso se consideriamo che la fenomenologia fonetica del linguaggio dei Down non è stata mai affrontata nella letteratura italiana. La nostra ricerca intende pertanto perseguire tale direzione di indagine, senza per questo avere pretesa di esaustività. Il carattere innovativo dello studio risiede nella metodologia adottata: la raccolta dei risultati è, per la prima volta in Italia, incentrata su una verifica di natura spettro-acustica. I dati discussi hanno carattere introduttivo; la ricerca, in corso di svolgimento, è stata di recente estesa.
I locutori presi in esame sono quattro, tre maschi e una femmina, di età compresa tra sei e dieci anni. Si tratta di bambini Down nati e vissuti in Calabria che frequentano la scuola dell'obbligo e settimanalmente si incontrano nella sede A.I.P.D. (*Associazione Italiana Persone Down*) di Vibo Valentia. Dopo un periodo di frequentazione, i soggetti sono stati audioregistrati, previo esplicito

consenso dei familiari e dei responsabili della struttura ospite. Il campione linguistico non è stato elicitato spontaneamente, bensì raccolto in un contesto verbale guidato; consapevoli dei limiti imposti dalla procedura metodologica adottata, si è ritenuto necessario, almeno in prima istanza, disporre di un controllo fine della situazione elocutiva[1].

Il materiale sottoposto all'attenzione dei soggetti è stato estratto da un *Test di Valutazione del Linguaggio* (TVL) elaborato da Cianchetti and Fancello (1997). Il test è costituito da cinque parti, per la precisione: 1) comprensione di parole e di frasi; 2) denominazione; 3) denominazione parole-figure; 4) ripetizione di frasi; 5) produzione spontanea su tema. In questa sede, considerato il carattere preliminare della ricerca, saranno discussi solo i dati relativi alle sezioni 3 e 4.

La terza sezione comprende ottantotto immagini denotanti oggetti di uso comune e familiare. Ai locutori è stato chiesto di denominare gli stimoli visivi di volta in volta presentati. La scelta delle immagini non è casuale, ma calibrata in modo da indurre l'elicitazione dei fonemi consonantici, scempi e geminati dell'italiano, in posizione iniziale e mediana di parola. La valutazione e il computo percentuale dei processi allofonici rinvenuti per ciascun fonema sono stati effettuati mediante l'ascolto del materiale e un'adeguata analisi spettrografica.

La quarta sezione consta da un Test di Ripetizione formato da quindici frasi aventi lunghezza e complessità sintattica progrediente. Si tratta di uno strumento volto a valutare il raggiungimento di alcune abilità linguistiche, in *primis* morfosintattiche, generalmente somministrato in bambini di età prescolare o in pazienti con ritardi linguistici (cfr. Devescovi and Caselli, 2001). Il test è finalizzato all'elicitazione di una frase precostituita formata da parole ad alta frequenza d'uso, la quale è indotta tramite imitazione verbale: il medesimo stimolo verbale è infatti realizzato poco prima dallo sperimentatore. La prova è stata eseguita singolarmente per ciascun soggetto in un ambiente sufficientemente silenzioso per non compromettere la qualità sonora del segnale audio.

[1] La raccolta del materiale audio è stata effettuata in collaborazione con Anna Cuccione nel 2008; si coglie l'occasione per ringraziare pubblicamente quanti hanno consentito il compimento di questa ricerca.

3 Analisi fonetica

Le registrazioni sono state trascritte foneticamente mediante l'alfabeto IPA; indi si è proceduto con il campionamento digitale del materiale audio (16000 Hz, 16 bits). L'analisi ha previsto l'ascolto reiterato del campione e l'analisi della struttura spettrografica; quest'ultima è stata compiuta mediante il *software Praat* (ver. 5.06.16). Sul versante fonologico, l'attenzione è stata diretta solo alla produzione delle parole isolate e al solo sistema consonantico. Inizialmente, si è proceduto con l'ispezione delle caratteristiche spettro-acustiche di tutti i fonemi consonantici occorrenti in contesto iniziale di parola ed intervocalico (cfr. § 4); successivamente, sono stati esaminati i fonemi geminati (cfr. § 5) ed infine i gruppi consonantici (cfr. 6). L'analisi ha permesso di computare tutti i casi di omissione e di sostituzione e di ottenere un quadro dettagliato dell'incidenza percentuale dei fenomeni osservati; sul piano qualitativo, l'indagine spettrografica ha portato alla definizione dei processi di sostituzione allofonica rinvenuti. La metodologia adottata è stata progressiva, dalla valutazione generale dei dati empirici si è passati, restringendo di volta in volta il campo di osservazione, all'analisi delle singole classi consonantiche.

Preliminarmente, occorre osservare che il linguaggio dei soggetti Down presenta una compromissione articolatoria che intacca, modificandolo, l'intero piano segmentale, sia vocalico sia consonantico. Parimenti coinvolto è anche il livello prosodico: struttura ritmica ed intonativa mostrano diversi tratti devianti; non di meno i maggiori indici di fluenza risultano alterati mentre l'assetto vocale manifesta una diffusa nasalizzazione. Fatta questa premessa, non sorprende che diverse porzioni di parlato risultino completamente inintelligibili, un tratto quest'ultimo frequentemente citato nella letteratura sull'argomento (cfr. Kumin, 2006).

Il primo parametro considerato è stato il grado di correttezza consonantica (*Percentage of Consonants Correct*, PCC, cfr. Roberts *et al*, 2005), un indice relativo al rapporto tra il numero delle consonanti correttamente realizzate e il numero totale delle consonanti *target*. Nel complesso, le consonanti conformi al bersaglio fonologico sono pari al 59%, i fonemi sostituiti o omessi rappresentano invece il 41%. Per una migliore lettura dei dati, i fonemi, ripartiti per classe consonantica, sono stati considerati nella loro duplice ricorrenza contestuale, iniziale e intervocalica. Nella produzione verbale dei soggetti Down,

sostituzione e cancellazione non hanno lo stesso peso, le sostituzioni sono di gran lunga più frequenti rispetto alle omissioni, come è evidente anche dalla rappresentazione istografica riportata nella fig. 1.

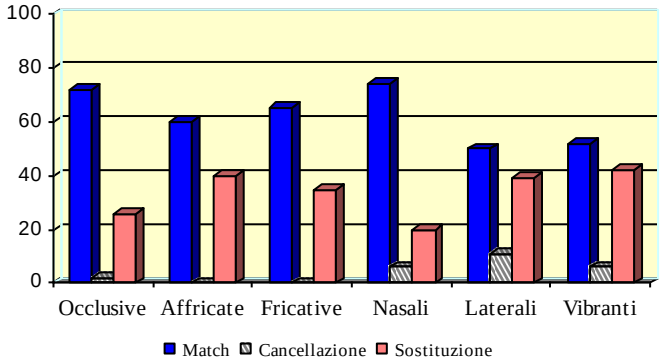


Figura 1: valori percentuali di corretta realizzazione (*Match*), Omissione e Cancellazione dei fonemi indicati

Nel dettaglio, il 31% delle consonanti riproduce una variante allofonica, significativamente inferiore è il tasso dei fonemi cancellati, 4%. La posizione occupata dal fonema all'interno del lessema influenza i due processi rilevati; in contesto iniziale le sostituzioni consonantiche sono maggiori (36%) rispetto a quanto avviene in posizione intervocalica, 25%. Per converso, la cancellazione del segmento ricorre più spesso nel contesto mediano (6%) che non in quello iniziale (2%). Sostituzione ed omissione inoltre non colpiscono in misura paritaria tutti i fonemi; come emerge dallo scorporo dei dati per classe consonantica (cfr. Fig. 2).

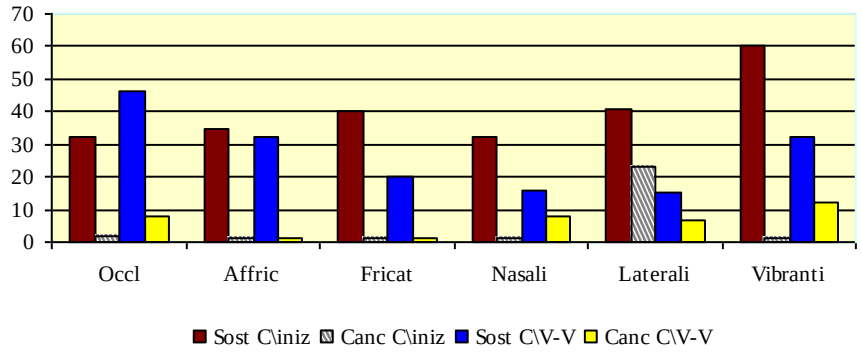


Figura 2: valori percentuali di sostituzione e di cancellazione in posizione iniziale (Sost. C\iniz, Canc. C\iniz) e intervocalica (Sost. C\V-V, Canc. C\V-V)

La sostituzione interessa tutti i modi consonantici, in special modo vibranti e laterali, è inoltre più frequente in posizione iniziale che non in contesto mediano; in controtendenza è tuttavia il comportamento delle occlusive, in special modo di quelle sonore, per le quali si rileva scarsa stabilità intervocalica (cfr. *ultra*). La cancellazione invece non coinvolge mai affricate e fricative, gli altri modi consonantici sembrano risentire del condizionamento posizionale: ad eccezione dei fonemi laterali, omessi in ambedue i contesti segmentali, il dileguo dei fonemi occlusivi, nasali e vibranti avviene unicamente in contesto intervocalico. In tab. 1 si riporta l'indice percentuale di ciascun processo rinvenuto ripartito per classe consonantica. La fenomenologia osservata sarà discussa singolarmente per modo consonantico.

	Occlusive	Affricate	Fricative	Nasali	Laterali	Vibranti
Sonorizzazione	18	5	9	--	--	--
Assordimento	7	--	4	--	--	--
Fricativizzazione	25	53	--	4	--	--
Affricazione	3	--	21	--	--	--
Nasalizzazione	2	--	--	--	5	23
Approssimantiz	27	--	26	--	10	6
Allungamento	5	26	26	36	15	6
Stopping	--	16	14	12	--	--
Lateralizzaz	7	--	--	16	--	--
Rotacismo	--	--	--	--	--	41
Velarizzaz	--	--	--	--	20	--
Anteriorizz	--	--	--	12	15	6
Omissione	6	--	--	20	35	18

Tabella 1: valori percentuali relativi ai processi allofonici osservati

4 La variazione allofonica

I fonemi occlusivi: La produzione delle occlusive è variabile, sebbene il luogo di articolazione del fonema *target* sia per lo più rispettato. I segmenti sordi sono realizzati fricativi o approssimanti, *dito* > [ˈdiːθo], le varianti allofoniche per le occlusive sonore, più varie e numerose, sono nuovamente quelle fricative, *baffi* > [ˈvafːi], approssimanti, *gara* > [ˈwaːra], ma si osserva anche qualche caso di nasalizzazione, *denti* [nɛ̃nti] o di lateralizzazione, *nido* > [ˈniːlo], due processi questi ultimi che denotano una mancata sincronizzazione dei gesti articolatori.

Per quanto riguarda la sonorità, la serie dei fonemi sordi /p t k/ subisce nel 18% un processo di sonorizzazione, *candela* > [gãnˈdɛːla], meno frequente è il processo inverso, l'assordimento di /b d g/ si rinviene infatti con una frequenza pari al 7%, *delfino* [tɘˈfːiːno].

Ciò che emerge è un diffuso indebolimento articolatorio, i soggetti Down approssimano il gesto di occlusione, senza però pervenire ad una totale chiusura degli organi articolatori. Quanto detto trova ulteriore conferma nel fatto che anche le occlusive sorde "ben realizzate" sono per lo più aspirate. L'aspirazione di un'occlusiva è universalmente associata ad un precipuo parametro acustico, il *Voice Onset Time* (VOT), ossia l'intervallo temporale che si estende dall'esplosione dell'occlusiva all'inizio delle vibrazioni periodiche del fono successivo[2]. A livello spettrografico, esso corrisponde ad un rumore di frizione di intensità variabile la cui durata è notoriamente correlata al luogo di articolazione, essendo breve nelle occlusive anteriori, lunga in quelle posteriori. L'aspirazione è distinta dal tratto articolatorio [+ glottide allargata], nella produzione di un'occlusiva sorda aspirata, la glottide è infatti già aperta, determinando così una concomitante fuoriuscita di aria. Dall'ispezione spettrografica condotta sui Down si evince come i segmenti sordi prodotti mostrino un VOT particolarmente lungo; analiticamente, questo parametro raggiunge la seguente durata media: /p/= 46 ms (DS: 19), /t/= 55 ms

[2] Fra tutti, Lisker
Barkat-Defrasas (2005), Sorianello (1996; 2008).

(DS: 26), /k/= 71 ms (DS: 28). Le occlusive sonore non sembrano condividere tale tratto fonetico, ma è pur vero che nel nostro corpus la ricorrenza di /b d g/ è sporadica, visto che la maggior parte di tali fonemi subisce variazione allofonica.

Nei soggetti Down, la ricorrenza di aspirate sorde è verosimilmente da ascriversi ad un controllo motorio insufficiente; rilassamento dei muscoli oro-facciali, diffusa ipotonia glottidale e respirazione scoordinata causano una articolazione consonantica imperfetta, che nel caso delle occlusive si traduce come apertura anticipata e prolungata della glottide. L'aspirazione, di norma intesa come processo di rafforzamento articolatorio, nel parlato delle persone Down è invece plausibilmente da considerarsi come indebolimento. Pur tuttavia, questo processo non ci risulta essere stato mai riportato per la Sindrome di Down; in riferimento al *pathological speech*, l'aspirazione dei segmenti occlusivi costituisce invece un tratto costantemente rinvenuto nei soggetti non udenti. Marotta (2002: 114) ritiene che l'allungamento del VOT negli audiolesi si debba «alla scarsa coordinazione tra gesti articolatori coinvolti nella fase di transizione tra i due stati della glottide (da aperta a chiusa)»; tuttavia, nei non udenti il percetto dell'aspirazione, più che alla lunghezza del VOT, sarebbe da ascriversi alla sua intensità sonora.

I fonemi affricati: Questi fonemi sembrano risentire poco del condizionamento contestuale. La fenomenologia osservata è riconducibile essenzialmente a due processi: fricativizzazione e resa occlusiva, un processo anche detto, seguendo la terminologia inglese, *stopping*. In ambedue i casi si tratta di una semplificazione del gesto affricato, o meglio di una sua scissione, visto che gli allofoni realizzano solo la parte occlusiva (22%), ad es. *zanzara* > [danˈdzaːra], *zaino* > [ˈdaːino] oppure, più spesso (71%), la sola porzione fricativa, ad es. *zucchero* > [ˈsukːelo]. Meno frequente (17%) è la sonorizzazione di una affricata sorda, ad es. *fucile* > [fuˈdʒiːle]. La realizzazione intervocalica di /dʒ/, in linea con la pronuncia centro meridionale, è rafforzata; a questa pronuncia si deve il dato relativo all'allungamento del fonema (26%).

I fonemi fricativi: Questa classe di suoni appare piuttosto stabile, soprattutto in posizione mediana; i fenomeni più frequenti rilevati a carico del modo fricativo sono l'allungamento iniziale (26%), ad es.

fumo [ˈfːumo], lo *stopping* (13%), ad es. *sole* > [ˈtɔːle], la sonorizzazione (9%), *befana* > [βeˈvaːna]. La variante approssimante riguarda in primo luogo \v\ che diventa [ʋ], un processo peraltro comune anche in italiano (cfr. Rispoli and Savy, 1997).
La percentuale di affricazione (22%) è da ascriversi invece a /ʃ/, sovente realizzato come affricata /tʃ/, ad es. *scivolo* > [ˈtʃːiʋolo], *sciarpa* > [ˈtʃapːa]. Questo processo di rafforzamento articolatorio potrebbe apparire in controtendenza rispetto alla tendenza, finora documentata, verso l'indebolimento; in realtà è l'espressione di un processo di ottimizzazione fonetica: le fricative palatali sono infatti suoni meno frequenti dei corrispettivi affricati, quindi più marcati sul piano tipologico. Le affricate /tʃ/ e /dʒ/ possiedono una salienza percettiva maggiore di /ʃ/ e /ʒ/, quest'ultimo peraltro assente nell'inventario fonologico dell'italiano, come di altre lingue (cfr. Maddieson, 1984). A ciò va aggiunto il fatto che il comportamento di /tʃ/ e /dʒ/ è, per più fattori, approssimabile a quello delle occlusive, segmenti forti per definizione (cfr. Baroni, 1993).
La stabilità della classe fricativa è tuttavia solo apparente (cfr. già Timmins *et al*, 2007), l'analisi spettrografica dimostra come la distribuzione lungo l'asse delle frequenze del rumore fricativo sia altamente variabile. L'impressione che ne deriva è che molti fonemi fricativi, *in primis* /s/, siano anteriorizzati e prodotti con un maggiore contatto linguale, alla stregua di suoni laterali, motivo per cui la frizione è meno turbolenta. La sostituzione di /s/ con [ɬ], ovvero laterale fricativa sorda, è d'altro canto una tipica manifestazione del parlato patologico, denominata sigmatismo laterale; si tratta di un fenomeno che tradisce un inadeguato contatto articolatorio insieme ad una impropria fuoriuscita d'aria.
I fonemi nasali: Il processo più rilevante a carico di /m/ e /n/ è l'allungamento (36%), l'incremento temporale si manifesta nelle due posizioni considerate: iniziale, ad es. mano > [ˈmːaːno], ed intervocalica: *banana* > [paˈnːaːna]. Per gli stessi fonemi, frequente (20%) è anche il dileguo del fonema, *pane* > [ˈpaːe], con conseguente allungamento della vocale. Il fonema palatale /ɲ/ non è mai realizzato; ciò che si verifica è un processo di scissione articolatoria classificabile come *phonemic split*, motivo per cui il suono *target* è prodotto come

la sequenza di una nasale alveolare e di una semiconsonante palatale, ovvero [n+j], ad es. *ragno* > [ˈlaːnjo]. La semplificazione del gesto articolatorio determina l'anteriorizzazione del luogo di articolazione del fonema palatale, il tratto di palatalità, assente dalla nasale, è veicolato dalla presenza del *glide* [j].

La pronuncia dei suoni nasali, insieme a quella delle occlusive, sembra nel complesso meglio tollerata rispetto agli altri modi consonantici; il dato acquisizionale riportato per la Sindrome sembra avvalorare questo dato; i suoni che compaiono prima nei soggetti Down, già nello stadio prelinguistico, sono per l'appunto le occlusive, orali e nasali, seguite dai *glides*, mentre gli ultimi ad essere realizzati sono i segmenti liquidi, fricativi e affricati (cfr. Stoel-Gammon, 2001).

I fonemi laterali: Nel 25% delle sue ricorrenze, il fonema /l/ non è realizzato, il dileguo si verifica tanto in posizione mediana, *scala* > [ˈkaːa], quanto in posizione iniziale di parola, *luce* > [ˈuːtʃe]. In modo speculare, rileviamo anche processi inversi, dunque non solo omissione, ma anche allungamento del fonema (15%), *valigia* > [taˈlːidːʒa]. La presenza di un allofono velarizzato (20%) sembra invece essere favorita dal contesto, poiché presente prevalentemente prima di una vocale posteriore, ad es. *luna* > [ˈɬuːna] o, ove realizzato, anche in coda sillabica. Il fonema palatale /ʎ/ è per lo più anteriorizzato, la sua realizzazione fonetica più frequente è infatti [l], oppure [l+j].

I fonemi vibranti: Il modo di articolazione vibrante si dimostra altamente instabile, ad ulteriore conferma dell'estrema variabilità articolatoria che contraddistingue questa classe di suoni nelle lingue naturali. Il ventaglio delle varianti allofoniche rilevato è ampio: nel 41% dei casi il fonema /r/ è pronunciato [l], per cui, *rana* [ˈlaːna], *pera* [ˈpɛːla]; la difficoltà meccanica di pervenire ad una corretta pronuncia di /r/ è stata definita rotacismo, seguendo la terminologia impiegata in ambito logopedico. È rilevante notare che il fenomeno opposto, sebbene più frequente e naturale nelle lingue, non è mai attestato, ovvero /l/ > [r]. Non mancano i casi di allungamento iniziale e di approssimantizzazione (6%), quest'ultimo processo fa riferimento alla sostituzione del fonema /r/ con il *glide* [j]. L'articolazione della vibrante suscita non poche difficoltà, praticamente assente è la

pronuncia polivibrante, i segmenti che riproducono la forma fonologica realizzano in realtà una monovibrante, o anche un'approssimante, ovvero [ɾ]. Non è insolito riscontrare queste varianti anche nell'italiano parlato, specie negli stili diafasici più veloci e meno controllati; tuttavia, nei bambini Down esaminati, questa pronuncia è plausibilmente da ascriversi ad un precipuo impedimento articolatorio: sia l'innalzamento del dorso della lingua sia il contatto alveolare, movimenti implicati nella produzione di /r/, risultano problematici, per via della condizione di macroglossia.

Complessivamente, la sostituzione si prefigura come un macrofenomeno all'interno del quale ricadono processi allofonici di diversa natura. I risultati empirici mostrano la presenza di una certa instabilità articolatoria, ogni modo consonantico manifesta un'ampia gamma di realizzazioni fonetiche. L'esito equivale il più delle volte alla riduzione di un gesto articolatorio complesso, come ad esempio la resa approssimante di un'ostruente o la realizzazione fricativa di un'affricata; non manca tuttavia il rafforzamento del fonema *target*, come l'affricazione di un suono fricativo o l'allungamento di una consonante scempia. Meno attestate sono le sostituzioni che generano uno spostamento del luogo di articolazione, fondamentalmente rileviamo l'anteriorizzazione e la velarizzazione, relativi rispettivamente alla pronuncia dei fonemi alveopalatali /ʃ ɲ ʎ/ e /l/. Anche il tratto di sonorità dei fonemi bersaglio non è sempre rispettato, sonorizzazione o, al contrario assordimento, coinvolgono esclusivamente le ostruenti; il dato non sorprende visto che le consonanti sonoranti sono per definizione [+ sonoro].

Nel campione linguistico indagato, la cancellazione colpisce in modo maggioritario le consonanti sonoranti, coinvolge poco le occlusive e mai i fonemi fricativi o affricati. La probabilità che un fonema sia completamente eliminato dalla stringa fonica sembra dunque essere direttamente collegato al grado di forza consonantica, i suoni scarsamente sonori e quindi più forti (occlusivi, affricati e fricativi) sono più resistenti, pur subendo nel corpus diversi processi di indebolimento. Per contro, le consonanti sonoranti, nasali, laterali e vibranti, sono segmenti intrinsecamente più deboli, poiché la quantità di energia necessaria per la loro produzione è minore; di conseguenza, il dileguo totale, stadio ultimo di questa deriva di progressivo indebolimento, è una condizione altamente predisponente. La

posizione preferenziale del fenomeno di cancellazione è quella intervocalica, un aspetto quest'ultimo che denota anche l'azione di un certo condizionamento sillabico.

5 I fonemi geminati

I soggetti Down indagati non mostrano particolari difficoltà con la pronuncia dei fonemi geminati, i quali sono replicati fedelmente nel 76% dei casi. La tipologia del suono implicato tuttavia condiziona l'esito finale, le consonanti fricative non subiscono degeminazione; al contrario di laterali e vibranti, si veda quanto riportato in tab. 2.

Modo Articolatorio	/CC/ > [Cː] %	Durata media ms	/CC/ > [C] %	Durata media ms
Occlusive	81	185 (57)	19	62 (6)
Affricate	69	178 (17)	31	87 (22)
Fricative	100	131 (33)	--	--
Nasali	90	150 (20)	10	86 (18)
Laterali	38	139 (7)	62	76 (11)
Vibranti	60	82 (19)	40	35 (12)

Tabella 2: valori percentuali, durata media e, tra parentesi, Deviazione Standard delle varianti geminate/CC/ > [Cː] e degeminate /CC/ > [C]

Alle geminate primarie si affiancano quelle secondarie determinate dall'allungamento spontaneo dei fonemi /b/ e /dʒ/; il modello di pronuncia centro-meridionale costituisce, per i giovani Down esaminati, un riferimento costante interiorizzato intriso anche di tratti dialettali. La buona padronanza mostrata dai Down nella produzione dei fonemi geminati può essere letta alla luce dell'alta salienza percettiva detenuta dai suoni lunghi. Il tratto fonologico dell'allungamento sembra inoltre correlarsi positivamente alla forza consonantica dei suoni medesimi; le consonanti liquide, suoni deboli e intrinsecamente più brevi, sono di frequente degeminate, laddove i suoni forti risultano, ancora una volta, più stabili anche sul versante temporale. Relativamente alle affricate, lo scempiamento è da considerarsi come la semplificazione di un gesto articolatorio di per sé

già complesso. Talora, il tratto di lunghezza del fonema affricato è foneticamente salvaguardato, a discapito tuttavia dell'identità del suono, quest'ultimo restituito da una variante fricativa, ad es. *tazza* > ['tasːa] o da un'occlusiva, *doccia* > ['dokːja].

L'intera categoria dei suoni [+ continuo] è interessato, con percentuali che oscillano tra il 6% e il 36%, da un allungamento secondario che si manifesta il più delle volte in contesto iniziale assoluto di parola. Tale allungamento temporale, a nostro avviso, non è da intendersi come un processo di geminazione, quanto piuttosto come una difficoltà di esecuzione motoria cui vanno incontro le persone con Sindrome di Down. L'eloquio dei soggetti indagati è molto spesso rallentato, l'allungamento iniziale dei suoni continui è da inquadrarsi in questo generale rallentamento del gesto articolatorio ed è pertanto classificabile come una disfluenza.

6 I nessi consonantici

L'analisi fonologica dimostra che il 41% di tutti i fonemi consonantici iniziali e intervocalici non riproduce il *target* articolatorio della voce lessicale somministrata. La situazione appare ancora più complessa se, accanto a questi dati, consideriamo anche la produzione dei gruppi consonantici. L'ipotesi di partenza è che la pronuncia sequenziale di più suoni consonantici implichi nelle persone Down un notevole sforzo articolatorio. I nessi consonantici sono stati analizzati spettograficamente. I processi osservati in questo specifico contesto sono diversi, per la precisione: a) totale cancellazione di un fonema, es. *freddo*> ['fɛdːo]; b) attivazione di un fenomeno di assimilazione totale, es. *erba* > ['ɛbːa]; c) inserzione di una vocale epentetica, es. *blu* > [bə'lu]; d) sostituzione di una consonante del nesso con una vocale, es. *sveglia* > [i'vɛːlja].

La fenomenologia osservata è stata inoltre ancorata alla struttura sillabica, al fine di verificare quanto il processo di riduzione risultasse condizionato dalla natura tauto o eterosillabica del nesso medesimo. Tutti i gruppi consonantici sono stati pertanto considerati separatamente, a seconda se il segmento ridotto occupasse la posizione strutturale di Attacco o di Coda sillabica. I contesti rinvenuti possono essere così classificati:

Attacco di sillaba: a) Ostruente + Liquida, ad es. *-tr-*, *-pl-*, *-fr-* ecc.
Coda di sillaba: c) /m n/ + C, ad es. *-nt-*, *-mb-* ecc.
d) /l/ + C, ad es. *-lt-*, *-lm-* ecc.
e) /r/ + C, ad es. *-rf-*, *-rk-* ecc.

Complessivamente, la produzione di un gruppo consonantico è minoritaria, solo il 27% dei nessi tautosillabici e il 22% di quelli eterosillabici risulta ben realizzato, i rimanenti gruppi, rispettivamente 73% e 78%, subiscono una profonda alterazione articolatoria.
Per quanto concerne l'Attacco sillabico, il 78% delle riduzioni si ha nel contesto O+/r/, ad es. *treno* > [ˈtɛːno]. Quando l'Attacco è ridotto, il segmento più stabile è, in linea con le nostre attese, il primo della sequenza, per cui /C_1C_2/ > [C_1]; nell'Attacco complesso, i segmenti in seconda o terza posizione, dotati di forza consonantica decrescente, sono infatti più esposti all'ipoarticolazione; i segmenti più propensi alla riduzione sono le consonanti liquide. Poco frequente è invece l'inserimento di una vocale epentetica, di norma uno *schwa* (10%), ad es. libro > [ˈliːbəro], ancora meno la sostituzione di una consonante dell'Attacco con una vocale (6%).
Il quadro fenomenologico relativo alla Coda sillabica è nettamente semplificato, l'unico processo rinvenuto è infatti la cancellazione. Le consonanti più frequentemente omesse sono /r/ (92%) e /l/ (84%), ad es. *farfalla* > [faˈfaːla], *altalena* > [ətʰæˈlɛːna]. Minore è la riduzione cui incorrono i fonemi nasali (56%), ad es. *campana* > [tʃiˈpãːna]. I fenomeni di assimilazione sono ben rappresentati; quasi sempre (88%) si tratta di un processo regressivo totale, per cui /$C_1.C_2$/ > [C_2ː], ad es. *delfino* > [təˈfːiːno], *sciarpa* >[ˈtʃapːa]. Le assimilazioni progressive rappresentano solo il 17%, ad es. *bimbo* > [ˈtimːo]. La predominanza di fenomeni regressivi ha una propria motivazione; il processo soddisfa infatti un principio di salienza fonetica: nella sequenza /$C_1.C_2$/, il segmento in Attacco, ossia C_2, è dominante, mostrando maggiore udibilità acustica e percettiva rispetto al segmento in Coda, cioè C_1 (cfr. Ohala, 1990). La spiccata tendenza verso l'assimilazione regressiva, chiaro fenomeno di *Coda weakening*, conferma lo statuto debole della Coda, una posizione strutturale universalmente meno solida rispetto all'Attacco.

Il comportamento sillabico di /s/, notoriamente problematico, merita di essere discusso a parte; numerose evidenze empiriche provano che lo statuto di /s/ preconsonantica sia eterosillabico, per cui si avrà *pas.ta*, *as.pro*, *cas.co*. Il quadro teorico è tuttavia complesso se consideriamo che le stesse sequenze possono comparire anche in posizione iniziale assoluta di parola, ad es. *stare*, *spronare*, *scoprire*. La soluzione fonologica di norma adottata è che /s/ sia in tale contesto dotata di natura extrametrica (cfr. Marotta, 1995; Bertinetto, 1999).
Nel corso della nostra analisi, incentrata su lessemi isolati, ci siamo imbattuti in tale contesto cruciale. L'ambiguità è stata fronteggiata e metodologicamente raggirata considerando le sequenze /s/+C ricorrenti all'inizio di parola tautosillabiche, e quelle ricorrenti all'interno eterosillabiche (già Chierchia, 1986). Nel campione analizzato, /s/ preconsonantica è omessa nel 65% dei casi se posta in Attacco, nel 50% se in Coda. Nel primo contesto, ovvero /s+C/ iniziale si osserva sempre la cancellazione della fricativa, ad es. *scala* > [ˈk^{h}aːa], mentre nel secondo, più che il dileguo, *castello* > [kaˈðɛːlo], si rinviene un assimilazione regressiva del nesso consonantico, ad es. *mosca* > [ˈmɔkː^{h}a].
Il campione di parlato patologico indagato evidenzia una significativa alterazione della struttura sillabica. Per i soggetti Down la pronuncia dei gruppi consonantici è per lo più evitata (già Rupela and Manjula, 2007): l'Attacco complesso è spesso semplificato mediante la cancellazione del segmento debole, quasi sempre una consonante liquida, o con l'epentesi di una vocale. Analogamente, le consonanti in Coda sillabica sono omesse o, in alternativa, assimilate al suono successivo. Il risultato è quasi sempre la riduzione del nesso consonantico e il ripristino della struttura sillabica non marcata CV.
Diverso è il quadro che scaturisce da un fenomeno assimilativo; in questo caso l'identità articolatoria dei due fonemi preserva la struttura sillabica, che mantiene la posizione della Coda, con conseguente chiusura della sillaba, per cui /C_1.C_2/ > [C_2ː]. Il risultato è un segmento lungo che legittima due posizioni sillabiche, per l'appunto la Coda e un Attacco semplice. La produzione di un segmento lungo è più facile rispetto ad un gruppo di consonanti, poiché implica l'esecuzione e il mantenimento temporale di un unico gesto articolatorio. Diversamente, la produzione di un nesso implica una coordinazione fine dei gesti articolatori implicati, una condizione

quest'ultima sfavorita dai tratti rilevati per questa Sindrome a carico dell'apparato fonatorio, oltre che da una scarsa coordinazione motoria.

7 Analisi morfologica

La morfologia è un'area linguistica problematica per le persone Down. Un tratto tipico della Sindrome documentato in più lingue concerne l'omissione dei morfemi grammaticali[3]. Per l'italiano, i dati sono stati raccolti raffrontando i campioni verbali prodotti da bambini Down e da gruppi di controllo costituiti da soggetti normodotati, con Sindrome di Williams o con Disturbo Specifico del Linguaggio (DSL)[4]. Dagli studi citati risulta che le categorie grammaticali per le quali si riscontra, tra i soggetti Down, il maggior numero di omissioni (mediamente 40/50%) sono le Preposizioni seguite dagli Articoli; Nomi e Modificatori sono al contrario gli elementi più rappresentati.
Non è tuttavia da escludere che un certo condizionamento possa essere esercitato anche dalla lingua materna dei soggetti Down; un indizio in questo senso sembra provenire da alcune ricerche condotte in contesto anglosassone. Jenkins (1993) osserva che le categorie meno impiegate da persone Down di lingua inglese, durante una prova di elicitazione guidata, siano in ordine di omissione: Ausiliari, Pronomi e Preposizioni. Il componente flessivo non è stato finora adeguatamente indagato relativamente ai soggetti trisomici di lingua italiana, diversamente da quanto avvenuto in altre lingue. O'Neill and Henry (2002) rilevano ad esempio come soggetti Down inglesi mostrino un diffuso *deficit* grammaticale che si manifesta anche in un uso errato dei morfemi legati, in particolar modo delle marche flessionali di tempo verbale (già Chapman *et al*, 1998).
Nel nostro studio, il comportamento morfologico dei soggetti Down è stato ricavato dall'elicitazione di un test di ripetizione di frasi (TRF). Nel campione esaminato non si rilevano errori di posizionamento delle forme lessicali e neppure errori formali derivanti da una errata

[3] Cfr. Chapman *et al* (1998); Eadie *et al* (2002); O'Neill and Henry (2002).
[4] Cfr. Fabbretti *et al* (1995); Vicari, Caselli & Tonucci (2000); Vicari *et al* (2002); Caselli *et al* (2006).

selezione delle marche morfologiche; la serie dei morfemi legati (ad es. categorie di genere, numero, tempo o aspetto verbale) è ben rispettata. I casi di sostituzione sono pochi, massiccio è però il tasso di omissione di forme lessicali. Le frasi prodotte integralmente costituiscono solo il 24% del campione. La lunghezza media delle frasi realizzate, a prescindere dal grado di completezza raggiunto, è stata computata assumendo quale riferimento sia il numero degli elementi lessicali sia il numero delle sillabe ivi contenuti[5], si veda in merito quanto riassunto in tab. 3. Significativa è la perdita di materiale lessicale che avviene nell'ultima parte dell'enunciato. C'è una correlazione direttamente proporzionale tra la lunghezza della frase e il suo grado di completezza; i locutori esaminati riproducono solo le frasi brevi, costituite in media da quattro o cinque elementi; gli enunciati progressivamente più lunghi sono quasi sempre incompleti; eloquente è anche il numero delle sillabe realizzate. Il deterioramento della Memoria a breve termine (MBT), tipico della Sindrome, rende difficoltosa la processazione di segnali verbali lunghi e l'elaborazione semantica del messaggio.

Correttezza frasi		Lunghezza frasi (parole)	Lunghezza frasi (sillabe)
Media	5	4	6
DS	2	2	3

Tabella 3: valori medi (X) e deviazione standard (DS) del numero delle frasi realizzate correttamente e della loro lunghezza in parole e in sillabe

La cancellazione degli elementi non è tuttavia casuale. La classe lessicale svolge un ruolo indiscusso, i lessemi più frequentemente omessi sono quelli funzionali, ben il 60%; per converso, le parole di contenuto, maggioritarie all'interno del campione verbale preso in esame, si prefigurano come elementi più robusti e meno inclini all'omissione (26%), come evidente dalla rappresentazione istografica riportata in fig. 3.

[5] Sono state conteggiate solo le sillabe realmente prodotte dai bambini Down e non quelle previste a livello ortografico. Nel computo, sono state considerate sia le parole di contenuto sia quelle funzionali.

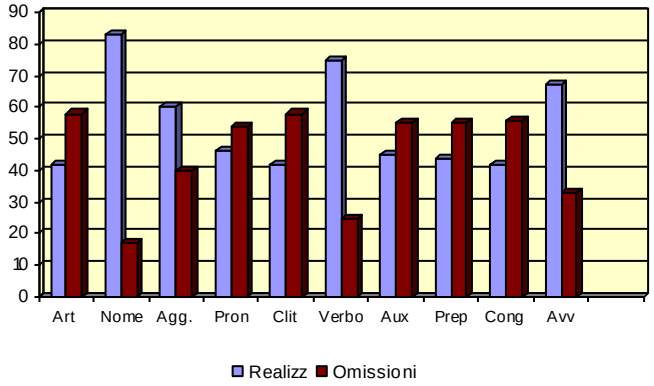


Figura 3: valori percentuali di corretta realizzazione (*Match*) e di Omissione delle diverse Classi Lessicali

I valori di omissione più alti riguardano gli Articoli determinativi, i Pronomi Clitici e le Preposizioni. Nelle frasi prodotte, gli Articoli sono frequentemente assenti (58%), all'interno della serie si notano alcune differenze: il determinante più realizzato è *la* (43%), *il* è prodotto solo nel 25% dei casi complessivi, inoltre nei contesti in cui è presente mostra sovente la cancellazione della consonante finale, ad es. *il cane* > [iˈkːʰăːne]. La presenza incisiva di *la* non sorprende, giacché realizza una condizione articolatoria ottimale: dal punto di vista sillabico, la struttura CV è di gran lunga più naturale rispetto alla sillaba chiusa, cioè lo schema sillabico di *il*, non a caso nella dinamica verbale sovente ridotto al solo componente nucleare. In aggiunta, la presenza della vocale /a/ concorre alla facilità articolatoria di *la*. Prevedibile è anche il comportamento dei Clitici, elementi scarsamente salienti sul versante fonetico, poiché brevi e, per definizione, deaccentati. Nel nostro campione, l'omissione di questa categoria raggiunge valori elevati, per la precisione il 58%. Lievemente più rappresentati sono i pronomi tonici (55%), tra le forme ricorrenti, *io* e *me* sono di norma prodotti; il carattere autoreferenziale conferisce a questi deittici personali minore vaghezza semantica.

Piuttosto alta è anche la percentuale di omissione delle Preposizioni (55%); gli elementi più realizzati sono *per*, *nel* e *con*, mentre risultano di norma omesse *a* e *in*. Da questi dati, ma il campione in realtà è

troppo esiguo per avanzare generalizzazioni, sembrerebbe che anche l'uso delle preposizioni sia vincolato in qualche misura al grado di pregnanza semantica; le preposizioni *a* e *in*, insieme a *di*, purtroppo non rappresentata nel corpus, possiedono il grado di astrattezza più elevato dell'intera categoria, e sono per questo anche definite "incolori". Analoga la tendenza al dileguo emersa per le Congiunzioni (56%), nel corpus compaiono solo elementi monosillabici (ad es. *e*, *se*), i quali condividono la stessa sorte di Articoli, Preposizioni e Clitici. Anche le forme verbali ausiliari esibiscono una diffusa fragilità, tra queste l'elemento più debole è la Copula. In questa classe si registra anche una sostituzione: l'enunciato *le rose hanno le spine* è infatti realizzato da uno dei soggetti indagati come [ˈrɔːsesɔ̃nəˈpːiːnə]. Nel complesso, la forma flessa *è* risulta quasi sempre omessa, al contrario della forma *hanno*; la maggiore lunghezza fonica contribuisce a rendere questo bisillabo più incisivo sul piano fonetico.

Le classi lessicali più resistenti sono quelle dei Nomi, Verbi e Avverbi, seguite dagli Aggettivi (cfr. fig. 3): alla presenza di questi elementi di contenuto, indispensabile per la comprensione degli enunciati prodotti, è affidata la riuscita, seppure parziale, del messaggio comunicativo. La forza di queste classi è data dalla confluenza sinergica di più fattori, priorità di comparsa nel processo di acquisizione della lingua materna e, relativamente ai Nomi e ai Verbi, universalità categoriale. Significativo è il comportamento degli Avverbi, nelle frasi ricorrono essenzialmente avverbi locativi e spaziali, riferimenti spazio-temporali legati alle esperienze sensoriali esperite dai soggetti fin dai primi mesi di vita e per questo precocemente acquisiti.

Dagli esiti empirici raccolti si desume che il numero delle categorie grammaticali non sia ridotto; tutte le categorie sono infatti impiegate, sebbene con percentuali differenziate, nessuna classe lessicale è totalmente omessa, anche se la ricorrenza di alcune è altamente instabile. La tendenza ad omettere certe classi lessicali durante l'esecuzione di un messaggio verbale non è un aspetto esclusivo della Sindrome di Down, il medesimo comportamento si registra in altre forme di parlato patologico, nelle prime fasi di acquisizione della lingua materna e finanche negli apprendenti una lingua seconda[6]. Le

[6] Questo parallelismo non è affatto nuovo; universalità dei comportamenti linguistici, in diacronia come pure in sincronia, ordine di precedenza e leggi di

cause come pure le modalità con le quali si manifesta tale fenomeno di erosione morfologica sono naturalmente diverse; pur tuttavia il raffronto dei dati può gettare luce sulla fenomenologia in discussione.
In riferimento al parlato patologico, anche i soggetti afasici, quelli affetti da DSL, da Sindrome di Williams e gli audiolesi tendono a sottoutilizzare alcune forme grammaticali; le categorie più fragili, pur rientrando nel novero degli elementi funzionali, collimano sostanzialmente con quelle rinvenute nelle persone Down. Da una ricerca condotta da Bottari *et al* (2001) su bambini con problemi linguistici emerge che Determinanti, Copula, Ausiliari, Clitici e Preposizioni, sono ridotti per numero; tuttavia nel soggetto afasico sono soprattutto le forme dei verbi ausiliari ad essere omesse (50% circa) laddove i bambini con DSL non producono i Determinanti nell'80% dei casi. Analogamente, anche il linguaggio verbale dei soggetti audiolesi presenta una morfologia lacunosa; nella lingua parlata le forme più inclini all'omissione sono quelle relative ad Articoli, Preposizioni, Ausiliari e Clitici, mentre nella lingua scritta si rilevano anche diverse riduzioni a carico dei morfemi legati (cfr. Aiello *et al*, 2001).
Tutto ciò potrebbe facilmente essere imputato ai deficit genetici o neurologici delle patologie citate, ma se allarghiamo il raggio di osservazione, si nota che una certa selezione morfologica contraddistingue il comportamento anche di bambini normodotati durante il periodo di acquisizione linguistica. I numerosi studi compiuti in questa direzione concordano su diversi punti[7], il percorso acquisizionale compiuto dai bambini nei primi tre anni di vita mostra un'alta omissione di morfemi funzionali. In aggiunta, l'acquisizione dei morfemi legati precede quella dei morfemi liberi, questi ultimi, generalmente omessi nei primi stadi linguistici, subiscono un progressivo incremento direttamente correlato con l'accrescere dell'età dei soggetti (cfr. Caprin and Guasti, 2009).
Non può passare inosservato il fatto che le medesime categorie funzionali compaiano tardi anche nel percorso acquisizionale di una

specularità, sono concetti più volte ribaditi e discussi da Jakobson secondo il quale: «la formazione del linguaggio infantile, la disgregazione del linguaggio afasico, la struttura e la modificazione delle lingue del mondo presentano una serie di leggi comuni di solidarietà» (Jakobson, 1971: 65).

[7] Cfr. Cipriani *et al* (1993); Pizzuto and Caselli (1992); Devescovi and Pizzuto (1995).

lingua seconda; a tale riguardo Giannini (2008) riporta come nell'interlingua di adulti anglofoni apprendenti l'italiano L2 il tasso di omissione dei pronomi clitici sia elevato; parimenti Calleri (2008) e Meini (2009), in riferimento all'italiano L2, mettono in luce l'alta instabilità detenuta dalle Preposizioni.
Quanto detto prova che le strutture morfologiche di una lingua vengano apprese per stadi, in modo graduale e inizialmente selettivo, ma non deviante. I comportamenti morfologici rilevabili nell'*iter* acquisizionale non sono mai accidentali. All'interno di una morfologia lacunosa, si ravvisa un *trend* regolare e per di più direzionale. Nei corpora acquisizionali (L1, L2) le parole di contenuto mostrano una fragilità minore rispetto alle parole funzionali, a riprova della priorità semantico-concettuale; una certa cronologia si ravvisa anche all'interno della sfera meramente grammaticale: i morfemi legati sono difatti utilizzati prima e meglio rispetto alle forme libere.
Degno di rilievo è tuttavia il fatto che le medesime tendenze, esasperate in termini percentuali, si osservino anche nel parlato patologico. La compromissione verbale rilevata nelle diverse patologie non intacca fondamentalmente gli schemi di acquisizione morfologica registrati nei soggetti normodotati. Così, l'eloquio delle persone Down, ma parimenti anche quello dei soggetti con DSL o con alcuni tipi di afasia, appare segnato da agrammatismo.
Cause esterne ed interne concorrono alla vulnerabilità delle categorie grammaticali[8]: fra le cause esterne è doveroso considerare la diversa natura degli elementi lessicali: le parole funzionali, sovente denominate anche elementi 'funtori', costituiscono dei sistemi chiusi, per numero e per funzioni, contrassegnati da rigida posizione contestuale. Queste forme, emblematicamente definite da Jakobson (1971) "accessori del linguaggio", sono inoltre prive di autonomia e poco incisive a livello fonetico, essendo brevi, spesso monosillabiche, e atone. Specularmente, i lessemi di contenuto sono elementi autonomi appartenenti a sistemi aperti; il loro grado di salienza percettiva è diverso, si tratta di parole più lunghe, generalmente toniche e sede virtuale di fenomeni di focalizzazione e di prominenza

[8] La fragilità del componente morfologico è d'altra parte testimoniata da una lunga serie di processi diacronici; all'origine di tale vulnerabilità ci sarebbe, secondo Simone (1988), l'azione di precipue proprietà semiotiche: esiguo corpo fonico, basso grado di iconicità e alto grado di articolazione.

prosodica. A tutto ciò va aggiunto, quale causa interna, come gli elementi funzionali, non a caso indicate come "parole vuote" abbiano un ridotto peso semantico. Bassa prominenza prosodica, impatto fonico ridotto e scarsa significanza finiscono per minare da un lato la salienza di questi elementi, e dall'altro la loro forza morfologica, rendendoli estremamente vulnerabili e, di conseguenza, fortemente predisposti alla cancellazione.
Ciò nonostante, a differenza di quanto avviene nell'acquisizione della lingua materna o di una lingua seconda, nei soggetti Down le abilità morfologiche non sembrano incrementabili; gli errori persistono nel tempo, contrassegnando il loro linguaggio anche nell'età adulta. La competenza linguistica delle persone Down mostra una sorta di cristallizzazione che blocca il processo acquisizionale ad uno stadio paragonabile a quella di bambini sani, come dimostra l'osservazione comparativa dei gruppi di controllo. Nella Sindrome di Down l'acquisizione del linguaggio avviene con un significativo rallentamento; è plausibile ipotizzare che al superamento del periodo critico, periodo compreso per la lingua materna tra i cinque e i dodici anni, i Down siano lontani dai livelli di competenza linguistica cui pervengono i loro coetanei normodotati. L'arresto del processo di sviluppo potrebbe spiegare perché la loro abilità linguistica non migliori nel tempo[9].

8 Riflessioni

La prima impressione linguistica che deriva dall'osservazione diretta dei locutori indagati è il basso grado di intelligibilità del loro eloquio; gran parte della produzione verbale spontanea dei soggetti non è agevolmente codificabile. Eppure il parlato Down mantiene, anche nei registri dialogici, una velocità elocutiva nettamente inferiore alla media. L'eloquio lento tuttavia non facilita la comprensione dei loro messaggi verbali, anzi in un certo senso sortisce l'effetto opposto, peggiorando la qualità percettiva. La lentezza elocutiva dei soggetti trisomici non scaturisce infatti da un meccanismo di iperarticolazione o di accuratezza linguistica, ma è il risultato di precisi disturbi, meglio noti con i termini di disfasìa e disartrìa, i quali finiscono con alterare i

[9] Per maggiori approfondimenti sull'argomento si rinvia a Lennenberg (1967) e Knudsen (2004).

principali indici di fluenza, determinando soventemente balbuzie, prolungamenti sonori impropri, veri sforzi articolatori. Lo scenario è ancora più aggravato se consideriamo che la fonazione dei soggetti Down è per lo più amodale: ipernasalità, bisbiglio, aspirazione e laringalizzazione sono i tratti più frequentemente rinvenuti nel corso della nostra ricerca (già Moura *et al*, 2009). A causa di un insufficiente controllo dei movimenti laringali, energia muscolare debole e problemi respiratori ricorrenti, la chiusura glottidale è spesso inconsistente; i cicli di adduzione e di abduzione delle pliche vocali sono irregolari ed incompleti, il risultato è un'impropria fuoriuscita di aria, percettivamente avvertita come *breathiness*.

Detto ciò, non è affatto sorprendente che il parlato delle persone Down esibisca una profonda compromissione fonologica che trova piena attuazione attraverso svariati processi di sostituzione allofonica. L'alterazione fonologica è ascrivibile essenzialmente a diversi ordini di fattori: anatomici, debolezza muscolare, pianificazione. Consideriamo innanzitutto la conformazione anatomica del loro apparato fonatorio; alcuni tratti come cavità orale piccola, lingua più larga della media, palato alto e stretto, costituiscono un notevole impedimento per l'articolazione di molti fonemi. L'iperlassità dei muscoli oro-facciali, laringali e della massa linguale concorre a peggiorare tale condizione, determinando un'imprecisione dei contatti articolatori e una precipua impressione acustica. A ciò si aggiunge una coordinazione motoria deficitaria la quale rende problematica la pianificazione e l'esecuzione di movimenti muscolari fini, come quelli verbali (cfr. anche Hamilton, 1993). Né va ignorato il ruolo giocato da una ridotta Memoria a breve termine: tanto la struttura come pure la lunghezza degli enunciati prodotti compromette, non solo la riuscita dei compiti di imitazione verbale, ma anche la pianificazione discorsiva spontanea. Il *deficit* a carico della memoria ha un effetto diretto anche sulla struttura morfologica della produzione verbale dei Down. Al fine di non sovraccaricare il peso mnemonico, di per sé già inficiato dalla Sindrome, le strutture frasali sono semplificate e alcune classi di parole sottoimpiegate. Si tratta di un meccanismo di economia azionato per fronteggiare uno *span* di memoria limitato che sortisce, quale effetto immediato, un parlato discontinuo costituito dalla sequenza di brevi blocchi semantici, privi il più delle volte di raccordo grammaticale, come dimostra la frequente omissione dei morfemi funzionali.

L'azione di questi agenti esterni non è tuttavia bastante a spiegare il comportamento verbale dei Down, né a cogliere la relazione profonda che si attua tra le forme di linguaggio disturbato e i normali processi di acquisizione linguistica. La soppressione delle categorie grammaticali, affatto caotica, si conforma infatti a principi di natura universale. Ciò che viene trattenuto possiede salienza percettiva, è questo il motivo per cui unità fonologiche brevi e atone, la cui processazione acustica è resa ancora più difficile in presenza di ipoacusia, tendono alla cancellazione. Le classi morfologiche più frequentemente conservate, tipicamente Nomi e Verbi, sono, oltre che salienti, anche naturali e poco marcate. A livello tipologico, Nomi e Verbi sono categorie cardinali universali presenti in tutte le lingue del mondo[10], ovvero classi prototipiche e per questo dotate di maggiore forza morfologica e robustezza semantica.
Anche in un contesto verbale "patologico" è possibile intravedere dunque una sorta di gerarchia delle unità linguistiche, tanto sul piano fonetico quanto su quello morfologico. Ad esempio, le vocali, suoni primitivi e acusticamente salienti, sono più stabili rispetto alle consonanti; parimenti, Nomi e Verbi, per i motivi suesposti, mostrano una solidità più elevata rispetto ad Articoli, Preposizioni o Clitici. Le unità linguistiche, gerarchizzate, sembrano quindi disporsi lungo un *continuum* ai cui estremi si pongono da un lato gli elementi forti e dall'altro le classi deboli con varie possibilità intermedie. I fenomeni osservati nei soggetti Down rafforzano l'ipotesi che acquisizione linguistica e disturbi del linguaggio vadano inquadrati e opportunamente interpretati come proiezioni superficiali di un'unica facoltà mentale: quella del linguaggio.

Riferimenti bibliografici

Ajello, R., G. Marotta, L. Mazzoni and F. Nicolai (2001) Morphosyntactic fragility in the spoken and written Italian of the deaf, in E. Fava (ed.) *Clinical linguistics. Theory and applications*

[10] Tale condizione non è vera per tutte le classi lessicali; diverse lingue mancano ad esempio dell'articolo, tra queste va menzionato almeno il latino; lo stesso dicasi per la categoria degli avverbi, anch'essa priva di carattere universale (cfr. Croft, 1991; Givón, 2001).

in speech pathology and therapy, Amsterdam-Philadelphia: John Benjamins, pp. 49-74.
Baroni, M. (1993) 'Teorie della sottospecificazione e restrizione sulle code consonantiche'. *Rivista di Grammatica Generativa* 18, pp. 3-59.
Bertinetto, P.M. (1999) La sillabazione dei nessi /sC/ in italiano: un'eccezione alla tendenza universale?, in P. Benincà, A. Mioni and L. Vanelli (eds) *Fonologia e morfologia dell'italiano e dei dialetti d'Italia*, Atti del XXXI Congresso della Società di Linguistica Italiana, Roma: Bulzoni, pp. 71-96.
Bottari, P., P. Cipriani, A. M. Chilosi and L. Pfanner (2001) 'The Italian determiner system in normal acquisition, specific language impairment, and childhood aphasia'. *Brain and Language* 77, pp. 283-93.
Calleri, D. (2008) Le preposizioni italiane: una categoria debole, in G. Bernini, L. Spreafico and A. Valentini (eds) *Competenze lessicali e discorsive nell'acquisizione di lingue seconde*, Perugia: Guerra Edizioni, pp. 37-52.
Caprin, C. and M.T. Guasti (2009) 'The acquisition of morphosyntax in Italian: a cross-sectional study'. *Applied psycholinguistics* 30, pp. 23-52.
Caselli, M.C., L. Monaco, M. Trasciani and S. Vicari (2006) 'La capacità di linguaggio in bambini con sindrome di Down e con Disturbo specifico del linguaggio'. *Psicologia clinica dello sviluppo* 3, pp. 473-91.
Chapman, R., H. Seung, S. Schwartz and E. Kay-Raining Bird (1998) 'Language skills of children and adolescents with Down Syndrome, II Production'. *Journal of speech, language and hearing research* 41, pp. 861-73.
Chierchia, G. (1986) 'Length, syllabification and the phonological cycle in Italian'. *Journal of Italian linguistics* 8, pp. 5-34.
Cho, T. and P. Ladefoged (1999) 'Variation and universal in VOT: Evidence from 18 languages'. *Journal of Phonetics* 27, pp. 207-29.
Cianchetti, C. and S. Fancello (1997) *TVL\ Test di valutazione del linguaggio* (30-71 mesi), Trento: Erikson.
Cipriani, P., A.M. Chilosi, P. Bottari and L. Pfanner (1993) *L'acquisizione della morfosintassi in italiano: fasi e processi*. Padova: Unipress.

Croft, W. (1991) *Syntactic categories and grammatical relations. The cognitive organization of information*. Chicago: University of Chicago Press.

Devescovi, A. and E. Pizzuto (1995) Lo sviluppo grammaticale in G. Sabbadini (ed.) *Manuale di neuropsicologia dell'età evolutiva*, Bologna: Zanichelli, pp. 260-85.

Devescovi, A. and M.C. Caselli (2001) 'Una prova di ripetizione di frasi per la valutazione del primo sviluppo grammaticale'. *Psicologia clinica dello sviluppo* 5, pp. 341-64.

Eadie, P., M. Fey, J. Douglas and C. Parsons (2002) 'Profiles of grammatical morphology and sentences imitation in children with specific language impairment and Down Syndrome'. *Journal of speech, language and hearing research* 45, pp. 720-32.

Fabbretti, D., E. Pizzuto, S. Vicari and V. Volterra (1995) L'acquisizione del linguaggio: caratteristiche della produzione verbale, in A. Contardi and S. Vicari (eds) *Le persone Down*, Milano: Franco Angeli, pp. 49-67.

Fowler, A.E. (1990) Language abilities in children with Down Syndrome: evidence for a specific delay, in D. Cicchetti, and M. Beeghley (eds) *Children with Down Syndrome: a developmental perspective*, Cambridge: Cambridge University Press, pp. 302-28.

Giannini, S. (2008) Proprietà formali e distribuzionali dei clitici in italiano L2, in R. Lazzeroni, E. Banfi, G. Bernini, M. Chini and G. Marotta (eds) *Diacronica et synchronica. Studi in onore di Anna Giacalone Ramat*, Pisa: ETS, pp. 231-54.

Gìvon, T. (2001) *Syntax. An introduction*. Amsterdam-Philadelphia: John Benjamins.

Hamilton, C. (1993) 'Investigation of the articulatory patterns of young adults with Down's syndrome using electropalatography'. *Down's syndrome research and practice* 1, pp. 15-28.

Jakobson, R. (1971)[2] *Il farsi e il disfarsi del linguaggio. Linguaggio infantile e afasia*. Torino: Einaudi.

Jenkins, C. (1993) 'Expressive language delay in children with Down's Syndrome. A specific cause for concern'. *Down's syndrome research and practice* 1, pp. 10-4.

Jerningan, T. L. and U. Bellugi (1990) 'Anomalous brain morphology on magnetic resonance images in Williams syndrome and Down syndrome'. *Archives of neurology* 47, pp. 529-33.

Knudsen, E.I. (2004) 'Sensitive periods in the development of the brain and behaviour'. *Journal of Cognitive Neuroscience* 16, pp. 1412-25.

Kumin, L. (2006) 'Speech intelligibility and childhood verbal apraxia in children with Down Syndrome'. *Down syndrome research and practice* 10, pp. 10-22.

Lennenberg, E. (1967) *Biological foundations of language*. New York: J. Wiley&Sons.

Lisker, L. and A.S. Abramson (1964) 'A cross-language study of voicing in initial stops: acoustic measurement'. *Word* 20, pp. 384-422.

Lubec, G. (2002) 'The brain in Down syndrome (Trisomy 21)'. *Journal of neurology* 249, pp. 1347-56.

Maddieson, I. (1984) *Patterns of sounds*. Cambridge: Cambridge University Press.

Marotta, G. (1995) La sibilante preconsonantica in italiano: questioni teoriche ed analisi sperimentale, in R. Ajello and R. Sani (eds) *Scritti linguistici e filologici in onore di Tristano Bolelli*, Pisa: Pacini, pp. 393-437.

– (2002) Voice Onset Time: un confronto tra italiani udenti e non udenti, in M.E. Favilla (ed.) *Comunicazione e sordità*, Pisa: Pacini, pp. 101-17.

Meini, L. (2009) *Dimensione dello spazio nelle preposizioni. Uno studio empirico sull'italiano*. Pisa: Editore Plus.

Mori, L. and M. Barkat-Defradas (2005) 'Acoustic properties of *foreign* accent: VOT variation in Moroccan-accented Italian'. *Interspeech 2005*, Lisbona, pp. 2909-12.

Moura C., L. Cunha, H. Vilarinho, M. Cunha, D. Freitas, M. Palha, S. Pueschel and M. Pais-Clemente (2009) 'Voice parameters in children with Down Syndrome'. *Journal of voice* 22, pp. 34-42.

Nicolai, F. (2003) *Argomenti di neurolinguistica. Normalità e patologia nel linguaggio*. Pisa: Edizioni del Cerro.

Ohala, J. (1990) The phonetics and phonology of aspects of assimilation, in J. Kingston and N. Smith (eds) *Papers in laboratory phonology I. Between the grammar and he physics of speech*, Cambridge: Cambridge University Press, pp. 258-75.

O'Neill, M. and A. Henry (2002) 'The grammatical morpheme difficulty in Down's Syndrome'. *Belfast working papers in language and linguistics* 15, pp. 65-72.

Pizzuto, E. and M.C. Caselli (1992) 'The acquisition of Italian morphology. Implications for models of language development'. *Journal of child language* 19, pp. 491-557.

Rispoli, G. and R. Savy (1993) 'Alcune considerazioni spettrografiche sulla [v] italiana'. *Studi italiani di linguistica teorica e applicata* 24, pp. 439-44.

Roberts, J., S.H. Long, C. Malkin, E. Barnes, M. Skinner, E.A Hennon and K. Anderson (2005) 'A comparison of phonological skills of boys with fragile X syndrome and Down syndrome'. *Journal of speech, language and hearing research* 48, pp. 980-95.

Rupela, V. and R. Manjula (2007) 'Phonotactic patterns in the speech of children with Down syndrome'. *Clinical linguistics and phonetics* 21, pp. 605-22.

Simone R (1988), 'Fragilità della morfologia e contesti turbati', in Giacalone Ramat A. (ed.), *L'italiano tra le altre lingue: strategie di acquisizione*, Bologna, Il Mulino, pp. 91-113.

Sorianello, P. (1996) 'Indici fonetici delle occlusive sorde nel cosentino'. *Rivista Italiana di Dialettologia* 20, pp. 123-59.

– (2008) Persistenza di categorie fonologiche nell'italiano L2, in L. Costamagna and G. Marotta (eds) *Processi fonetici e categorie fonologiche nell'acquisizione dell'italiano*, Pisa: Pacini, pp. 85-110.

Stoel-Gammon, C. (2001) Down syndrome phonology: developmental patterns and intervention strategies. *Down syndrome research and practice* 7, pp. 93-100.

Timmins, C., W. Hardcastle, S. Wood, J. McCann and J. Wishart (2007) 'Variability in fricative production of young people with Down's Syndrome: an EPG analysis', in Trouvain J. and W. J. Barry, *Proceedings of the 16th International Congress of Phonetic Sciences*, Saarbrucken, Germany, pp. 1981-1984.

Vicari, S., M.C. Caselli and F. Tonucci (2000) 'Asyncrony of lexical and morphosyntactic development in children with Down syndrome'. *Neuropsychologia* 38, pp. 634-44.

Vicari, S., M.C. Caselli, C. Gagliardi, F. Tonucci and V. Volterra (2002) 'Language acquisition in special populations: a comparison between Down and William Syndromes'. *Neuropsychologia* 40, pp. 2461-70.

Wang, P.P., J.R. Hesselink, T.L. Jernigan, S. Doherty and U. Bellugi (1992) 'Callosal morphology concurs with neurobehaviour and

neuropathological findings in two neurodevelopmental syndromes'. *Archives of neurology* 49, pp. 407-11.

PARTE VII

STRUMENTI E TECNICHE PER LO STUDIO DELLA COMUNICAZIONE PARLATA

Il processo di Erba tra Analisi Transazionale e *Counselling* giuridico

Riflessioni per una comunicazione efficace[*]

Monica Dell'Aglio – Maria Lasaponara
Unibas – Campus srl
mdellaglio@gmail.com; *maria.lasaponara@libero.it*

Abstract

The one who plays the "director" role in a face to face asymmetric communication, must be able to use all instruments to manage with effectiveness the communication rapport as well as his social role. The technical mastery of his work, the high language competence and the psychological opening to other people, supply him these instruments. All these skills are indispensable when he is a lawyer, a judge, a prosecutor or a police officer because of the hard, painful situation of the other.
This work analyzes the linguistic exchanges between prosecutors and Olindo and Rosa-Angela Romano spouses (inquired for the massacre of Erba) compared to the theoretical frames from Transitional Analysis and the Legal Counselling.
The use of these patterns evidences that the prosecutors were fully oriented towards their role. They were not able to create a rapport such as to enable communication between the "adult" alter ego state. So the Bazzis perceived an unequal strength distribution and adapted their expressive modality to the one of their prosecutors.
So the "director" role failed in leasing communication in an effective way.

1 Introduzione

Il presente lavoro si inserisce in un filone di studi sull'interazione faccia a faccia di tipo "asimmetrico"', in cui il potere comunicativo che gli interlocutori assumono all'interno della struttura partecipativa non risulta paritario, ma sbilanciato a favore di uno di essi a cui va il ruolo di regista e anche di "giudice". È quanto accade, ad esempio, tra insegnante-allievo, medico-paziente, avvocato-cliente, tra i partecipanti a un talk-show, a una tavola rotonda o a una selezione del personale (Bazzanella, 2008: 217-8; Orletti, 2000: 12-44; Petrillo, 1994). In questa sede osserveremo l'interazione all'interno di una

[*] Questo lavoro è frutto di una continua e comune elaborazione. Tuttavia, ai fini accademici, l'attribuzione dei paragrafi è la seguente: Monica Dell'Aglio: §§ 1, 2, 4, 4.1, 5; Maria Lasaponara: §§ 3,3. 1, 3.2.

comunicazione istituzionale che vede da un lato i Pubblici Ministeri e dall'altro gli indagati (Bellucci, 2002: 155-67). In particolare, cercheremo di capire se e in quale misura le figure "dominanti" siano consapevoli del fatto che le comunicazioni pubbliche, al pari di quelle personali, implicano valenze emotive, reazioni e resistenze la cui comprensione è necessaria ai fini di una gestione efficace non solo dell'evento comunicativo in sé, ma anche del proprio ruolo (Quadrio, 1993: 4). In ambito giudiziario, in cui l'interazione si configura come «luogo di costante trattazione e negoziazione» (Cavicchioli, 1997: 98), le figure istituzionali dovrebbero dunque mostrare una preparazione tecnica, una competenza linguistica – che troppo spesso risulta inadeguata e mascherata da un linguaggio burocratico che da molti è percepito come rinforzo della propria identità sociale (Bellucci, 2002: 148) – nonché psicologica dovendosi esse confrontare con interlocutori che chiamano in causa temi personali ed emotivamente significativi (Guida, 1993). Per evidenziare il possesso e la gestione di tali capacità, nonché la rilevanza comunicativa delle risposte fornite dagli inquisiti nel modificare la comunicazione degli inquisitori, abbiamo analizzato gli eventi linguistici alla luce del quadro teorico fornito dall'Analisi Transazionale (AT) e dal *Counselling* giuridico.

2 Materiali

La nostra analisi si fonda su diversi materiali relativi alla strage di Erba dell'11 dicembre 2006 per la quale sono stati prima indagati e poi condannati Olindo Romano e Rosa Angela Bazzi[1].
L'efferatezza dell'esecuzione ha creato intorno alla vicenda un clima di grande interesse, e i mezzi di comunicazione si sono avvalsi di qualsiasi canale per diffondere le notizie di cui entravano in possesso. Da subito infatti è stato possibile reperire on-line trascrizioni di dichiarazioni, interrogatori e intercettazioni ambientali, oltre che gli articoli pubblicati dalle testate giornalistiche nazionali (http://www.laprovinciadicomo.it/publisher/cronaca/section/; http://www.corriere.it/cronache/; http://www.repubblica.it/cronaca/indici/index.html; http://wilgreta.wordpress.com).

[1] I coniugi sono stati riconosciuti colpevoli dell'omicidio dei signori Raffaella Castagna, Youssef Mourzuk, Paola Galli e Valeria Cherubini, nonché del ferimento del signor Mario Frigerio.

Presso l'ufficio Rai "Teche Aperte" di Potenza abbiamo effettuato le prime ricognizioni del materiale che veniva messo in onda (TG2, Tg2Dossier, Un Giorno in Pretura, Porta a Porta, La Vita in Diretta, e molto altro). Parte di questo materiale è stato poi acquisito dalla Teca Rai di Roma poiché non era ancora disponibile on-line.
In questa sede, per motivi di spazio, faremo riferimento ai soli dati tratti dalle analisi effettuate sul verbale relativo all'interrogatorio del signor Olindo Romano avvenuto il 10 gennaio 2007 alle ore 14.45, nonché a quello della signora Rosa Angela Bazzi avvenuto nello stesso giorno alle ore 15.25[2]. I verbali hanno fornito dati sufficienti per inquadrare le dinamiche dell'interazione verbale e, a partire dai requisiti di efficacia e di efficienza comunicativa, per esemplificare la circolarità e il contenuto emotivo propri di ogni interazione.

3 Analisi Transazionale

L'Analisi Transazionale (AT), elaborata da Eric Berne in chiave psicoanalitica (Berne, 1971; 1979), è una teoria sistematica e coerente della personalità, che definisce rapporti intra- ed interpersonali alla luce di una particolare strutturazione del proprio essere.
La personalità, sempre secondo Berne, va studiata in riferimento a quattro diverse tipologie di analisi:

Analisi Strutturale: studio della personalità distinta in stati dell'Io;
Analisi delle Transazioni: studio delle modalità relazionali;
Analisi di Passatempi e Giochi: studio delle manovre sociali utilizzate dagli individui nelle relazioni e nei gruppi;
Analisi del Copione: studio dello schema di vita inconscio che limita e condiziona le modalità relazionali dell'individuo sulla base di programmi cognitivi, emotivi e comportamentali.

Partendo dagli stati dell'Io lo studioso riesce a spiegare tanto i disturbi psicologici, quanto il comportamento umano nella sua dimensione interpersonale. Ne deriva, dunque, un modello di terapia caratterizzato da una componente attivistica e da una razionale. La prima attribuisce

[2] Il materiale analizzato ci ha consentito di riflettere anche su molti altri aspetti – problemi di trascrizione; incongruenza tra i materiali audio e le trasposizioni cartacee; linguaggio burocratico; la rappresentazione dello straniero ecc. – che tratteremo in altra sede.

al cliente, o paziente, la capacità di esercitare un autocontrollo sul proprio comportamento e, di conseguenza, la gestione delle proprie azioni e delle motivazioni che ne sono alla base. La seconda, invece, oltre a richiedere al cliente un coinvolgimento attivo nel processo di scambio, riguarda le abilità che il terapeuta deve avere nell'esercitare un "controllo ragionato" sull'interazione. Pertanto, la comunicazione, frutto di un continuo confronto all'interno di una terapia non direttiva, risulta regolata da un sistema di autovalutazione che rende semplice la scoperta di elementi discordanti nella personalità (Moiso and Novellino, 1982: 74-5). Inoltre, la possibilità di isolare gli stati dell'Io consente di riconoscere la posizione esistenziale che l'altro riveste all'interno del circolo comunicativo (Harris, 2000: 32).
In questa sede ci soffermeremo solo sull'analisi della Struttura e delle Transazioni in quanto rappresentano non solo l'oggetto primario dell'AT, ma anche la chiave di lettura del nostro materiale.

3.1 Struttura e Funzione della personalità

Lo stato dell'Io viene definito di solito come un insieme di pensieri e sensazioni che si manifestano in modelli di comportamento corrispondenti (Berne, 1979: 18). A seconda delle relazioni che il soggetto costruisce, tali modelli tendono non solo a modificarne le espressioni verbali, paraverbali e non verbali, ma anche a spingerlo ad occupare posizioni esistenziali differenti (Harris, 2000). Questo è possibile perché ogni essere umano presenta tre diversi stati dell'Io, rappresentati graficamente dal modello **GAB**:

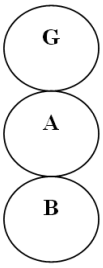


Figura 1: I tre stati dell'Io

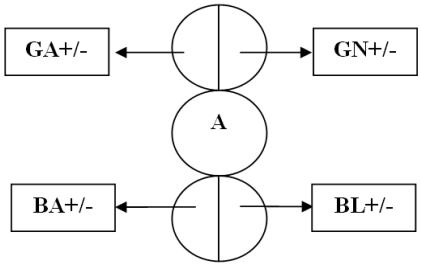


Figura 2: Le funzioni dei tre stati dell'Io

Il primo stato è quello dell'Io **Genitore** (**G**) che proviene dalle posizioni genitoriali e si manifesta attraverso emozioni, pensieri e

comportamenti appartenenti a tali figure di riferimento. Quando si trova in questo stato il soggetto si comporta, pensa, agisce e reagisce come faceva uno dei suoi genitori, assumendo il ruolo di "coscienza" nel rapporto con sé e con gli altri. Dal punto di vista funzionale questa sfera distingue due componenti: una genitorialità di tipo **Affettivo** (**GA**) che *insegna*, *consiglia*, *sostiene*, *aiuta*; e una genitorialità di tipo **Normativo** (**GN**) che *comanda*, *giudica*, *critica*, *controlla*. Entrambe le genitorialità possono essere positive o negative in relazione all'intenzione presente nel messaggio inviato. Nel caso di un **GN** o di un **GA** di tipo positivo, lo scambio sarà regolato da un intento volto alla valorizzazione dell'interlocutore, nonché alla costruzione di un rapporto solido e allo scambio "sano" tra personalità differenti. Al contrario, in una genitorialità di tipo negativo il messaggio sarà diretto ad annullare l'interlocutore e a sottometterlo.

Il secondo stato è espressione di razionalità. Si tratta dello stato dell'Io **Adulto** (**A**) attraverso cui il soggetto è in grado di giudicare e valutare l'ambiente circostante in base all'esperienza passata e al grado di maturità acquisito. La funzione di **A** è quella di mediare tra **G** e **B** in modo da rispondere agli stimoli esterni con il controllo della propria emotività.

Infine, il terzo stato si riferisce allo stato dell'Io **Bambino** (**B**) e rappresenta la parte più preziosa della personalità. Con esso il soggetto manifesta le sue volontà svincolate dal contesto di riferimento, e dà spazio alle passioni, ai desideri e ai sogni assumendo un comportamento libero dalle pressioni genitoriali e dal controllo dell'Adulto. Anche in (**B**) si distinguono due componenti principali: il **Bambino Libero** (**BL**), che adotta un comportamento svincolato dal contesto e dalle aspettative altrui, e il **Bambino Adattato** (**BA**), che al contrario modella il proprio comportamento in base alle esigenze e alle aspettative delle figure genitoriali. Anche in questi casi sono previste accezioni positive e negative a seconda che l'intenzione presente nel messaggio sia orientata alla costruzione o alla distruzione della relazione.

Dunque, nell'interazione tra due persone intervengono sei stati dell'Io, tre per ogni persona, e nello scambio tra gli stati è possibile riconoscere diverse tipologie di transazioni[3]. Nelle più semplici, che

[3] Per uno studio approfondito delle transazioni rimandiamo a Berne (1971).

sono di tipo *parallelo* e *complementare*, la comunicazione può durare all'infinito

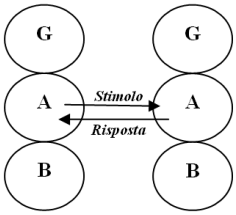


Figura 3: Transazione **AA-AA**

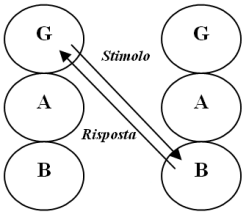


Figura 4: Transazione **GB-BG**

In Figura 3 è riportata la transazione che per l'AT è alla base di una "comunicazione efficace", vale a dire una transazione *parallela* e *complementare* tra due stati dell'Io Adulto.
Ci sono poi le transazioni di tipo *incrociato* che coinvolgono stati dell'Io differenti. La transazione, ad esempio, parte con uno stimolo dall'Adulto diretto all'Adulto, ma riceve risposta dallo stato dell'Io Genitore diretto al Bambino

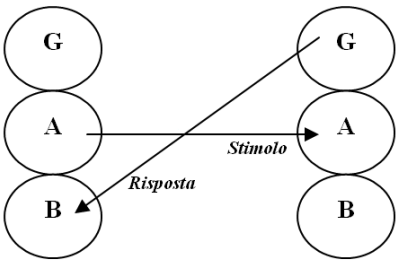


Figura 5: Transazione **AA-GB**

Tanto le transazioni *parallele*, quanto quelle *incrociate* risultano semplici dal momento che implicano solo il piano sociale. Al contrario, le transizioni a due livelli, definite *angolari* e *duplici*, rimandano ad un messaggio in cui è presente un contenuto manifesto, o sociale, ed un contenuto latente, o psicologico. Un esempio di transazione *angolare* è costituito da un messaggio manifesto che parte dallo stato dell'Io Adulto ed è rivolto allo stato dell'Io Adulto dell'interlocutore, con un livello psicologico che coinvolge invece lo stato dell'Io Bambino. Se la transazione ha successo, la risposta partirà proprio dallo stato dell'io Bambino (figura 6)

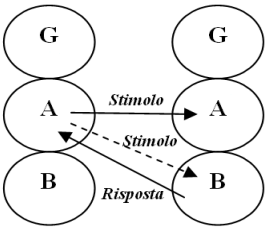


Figura 6: Transazione angolare

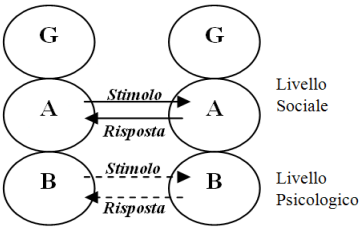


Figura 7: Transazione duplice

Nella transazione *duplice*, invece, il contenuto sociale è completamente differente dal contenuto psicologico. Ad esempio, può accadere che lo stimolo manifesto parta dall'Adulto e sia diretto all'Adulto, mentre quello latente va dal Bambino al Bambino (figura 7).

3.2 Analisi della Strage di Erba in chiave di AT

Nella prima fase del lavoro si è analizzato l'evento comunicativo intercorso tra i P.M. – Astori e Fadda – e il signor Olindo Romano. La lettura incrociata, in chiave analitico-transazionale, dei dati linguistici e dei dati psicologici ci ha permesso di individuare gli stati dominanti dell'Io e di riflettere sul legame tra questi e la riuscita o il fallimento di alcune mosse comunicative. Negli esempi il grassetto è stato aggiunto per richiamare l'attenzione sugli elementi ritenuti più significativi ai fini dell'analisi.

Esempio 1
[…]
P.M. Dott. ASTORI – Allora, è già generalizzato. **Sono le ore 14 e 45, è presente il difensore**, e questo **abbiamo già verbalizzato** tutto, benissimo. Lei come sempre **ha facoltà di non rispondere** se vuole alle nostre domande. **Intende rispondere?**
Indagato ROMANO OLINDO – **Sì, sì**.
[…]

Lo scambio esemplifica una **transazione parallela** che, coinvolgendo due stati dell'Io Adulto, risulta pienamente efficace. Il P.M. mette in atto un comportamento razionale di analisi ponendo una domanda diretta a cui l'interlocutore risponde in maniera pertinente. Entrambi si attengono ad uno schema del "qui e ora", annullando qualsiasi possibilità di incomprensione.

Mantenere costante lo scambio tra stati dell'Io Adulto risulta però molto difficile, soprattutto quando l'interazione è carica di forti contenuti emotivi:

Esempio 2
[…]
P.M. Dott. ASTORI – **Vuole rispondere**. **Senta signor Romano, dobbiamo riparlare un po' di quei fatti** anche perché come abbiamo appena comunicato anche al suo difensore, **all'interno della sua auto è stata ritrovata una traccia** che oltre ad essere del sangue della signora Cherubini, c'è anche quello del signor Frigerio e quindi **la sua posizione si è aggravata** ulteriormente. Quindi i primi esami del Dna oltre ad avere dato positività per la signora Cherubini, l'hanno dato anche per la signora…, per la signora Frigerio, l'hanno dato anche per il signor Frigerio. Quindi…
P.M. Dott. NALESSO – Allora, siccome lei sarà poi sentito dal G.I.P., **noi siamo venuti a chiederle ed a dirle** se non è il caso che, ragionevolmente, lei non ci dica che cosa è successo quella sera, perché **lei non è un bambino, noi non siamo bambini, non stiamo scherzando** e lei l'unica cosa che può fare per se stesso e per sua moglie, è di **dirci che cosa è successo** quella sera, **senza prenderci in giro**.
Indagato ROMANO OLINDO – **Io adesso tutto tutto non me lo ricordo proprio tutto tutto.**
[…]

La lettura congiunta delle mosse conversazionali dei due P.M. fornisce un esempio di **transazione angolare** riuscita. Facendo riferimento ai dati provenienti dall'esito delle indagini effettuate sull'autovettura dell'indagato, Astori sembrerebbe avviare una transazione parallela di tipo **AA-AA**. L'intervento del P.M. Nalesso ne chiarifica la natura angolare: questi infatti si rivolge all'Io Bambino del Signor Romano, che risponde con un linguaggio appartenente all'Io Bambino Adattato.
La comunicazione da qui in avanti procede attraverso lo scambio quasi costante tra due stati dell'Io predominanti. Lo stato dell'Io dei P.M. si mantiene sull'asse della genitorialità, e quello dell'indagato sull'Io Bambino, che non sempre però si configura come Adattato. Si consideri l'esempio seguente:

Esempio 3
[…]
P.M. Dott. NALESSO – Senta, se non ci vuole dire i particolari, a grandi linee **quello che si ricorda, quello che si sente**.
Indagato ROMANO OLINDO – **Adesso come adesso…**
P.M. Dott. FADDA – **Da dove vuole iniziare lei. Lei da dove vuole iniziare?**

Indagato ROMANO OLINDO – Io **stavo…, stavo pensando ad un'altra cosa adesso come adesso.** Questo può essere il minore dei mali, giusto? Sia per me che per mia moglie…
[…]

Il frammento presenta **transazioni complementari**. Lo stimolo dei P.M. è dettato dall'Io Genitore Normativo Positivo che si rivolge all'Io Bambino dell'indagato, ma questi mette in atto un comportamento da Bambino Libero svincolandosi sia dal contesto, sia dalla domanda che gli viene posta. Pur essendo complementari, queste transazioni non risultano efficaci, e quindi utili ai fini delle indagini, in quanto non riportano gli interlocutori sul piano del "qui e ora".
Nell'interrogatorio sono presenti anche casi di **transazione incrociata**:

Esempio 4
[…]
P.M. Dott. FADDA – **Dai signor Romano stavamo facendo**…, come le diceva prima il mio collega, dei **discorsi da persone adulte…**
Indagato ROMANO OLINDO – **Posso andare**, ascolti, **un attimo ancora da mia moglie e poi torno? Un secondo e basta?**
P.M. Dott. ASTORI – **No**.
P.M. Dott. NALESSO – No, adesso **basta non stiamo mica a giocare**.
[…]

Il signor Romano continua a ricoprire il ruolo di Bambino, e sembra che sia il P.M. Fadda a spingerlo in questa direzione, apparentemente in maniera involontaria. Le sue richieste infatti risultano non pertinenti alle domande che gli vengono poste. La transazione parallela da Genitore Normativo Positivo a Bambino Libero subisce un incrocio dovuto all'Adulto del P.M. Astori che interviene con un netto rifiuto. Il Dott. Nalesso ricade nella posizione del Genitore Normativo che cerca di mitigare il rifiuto del collega riportando l'attenzione dell'indagato al contesto comunicativo in atto.

4 *Counselling* e superamento dei conflitti

Il *counselling* è una disciplina giovane – nasce negli Stati Uniti negli '50 e arriva in Europa negli anni '70 – e viene impiegata sia come attività di orientamento, sia come strumento di supporto a servizi

sociali e di volontariato (Biggio: 2005, VII). A partire dai fondamenti di Pragmatica della Comunicazione, Analisi Transazionale e Programmazione Neuro Linguistica, il *counselling* forma figure professionali in grado di riconoscere e dominare le dinamiche innescate dal canale emotivo proprio e altrui (Vinella and Lopez, 2000) e di condurre il proprio cliente alla comprensione e alla risoluzione dei problemi (Biggio, 2005: 16-7, 79-80).
I settori che ricorrono al *counselling* come prassi operativa sono molteplici (Miglionico, 2000; Biggio, 2005); da qualche anno si sta promuovendo una sua applicazione anche in ambito giuridico[4]. In questo settore, l'applicazione del *counselling* risponde alla necessità di fornire una chiave di lettura della relazione umana svincolata, nei limiti del possibile, dalla materia legale e dalle norme che la regolano. Il *counsellor* quindi aiuta il cliente a superare il disagio emotivo che scaturisce dalla distribuzione impari di potere nella relazione, attraverso un'attività di esplorazione e intuizione emotivo-cognitiva. Metodologicamente, il *counsellor* opera facendo uso di tecniche di chiarificazione mirate a gestire la relazione diadica attraverso un atteggiamento prima maieutico e poi direttivo in grado di favorire lo scambio di informazioni, la comprensione dei messaggi e, cosa auspicabile, la libera confessione dei reati commessi.

4.1 Processo di Erba: riflessioni per un corretto intervento di *counselling*

Nella seconda parte del lavoro abbiamo riflettuto sulle principali fasi applicative dell'intervento di *counselling* (Biggio, 2005: 81-94; Di Fabio, 1999) traendo esemplificazioni dagli scambi conversazionali intercorsi tra i P.M. e la signora Rosa Angela Bazzi.
Nella prima fase del rapporto, detta di accoglienza, il *counsellor* deve innanzitutto creare un clima di fiducia mostrando una grande apertura verso il proprio cliente, in maniera che questi possa esprimere liberamente la sua versione dei fatti e, nel contempo, percepire che ci sia il desiderio di una collaborazione reciproca. In questo c'è un chiaro richiamo alla strategia di cortesia linguistica positiva, grazie alla quale si è anche in grado di salvaguardare la faccia negativa

[4] Si veda, tra i tanti, quello promosso dal C.IR.M. di Avellino (http://www.fiped.it/CORSI&P.htm).

dell'interlocutore. In pragmatica la faccia corrisponde all'immagine pubblica che ognuno ha di sé e che si aspetta venga riconosciuta dagli altri. Una persona ha una faccia negativa quando cerca autonomia e quindi rifiuta le imposizioni altrui; ne ha una positiva quando al contrario cerca l'approvazione degli altri, ovvero quando avverte il bisogno di sentirsi parte di un gruppo (Brown and Levinson, 1987: 61).
Chi riveste il ruolo di regista dell'evento comunicativo deve dunque conoscere questi aspetti, soprattutto se il contesto è caratterizzato da un'asimmetria istituzionalmente riconosciuta (Bellucci, 2002: 155).
Vediamo cosa avviene durante l'interrogatorio:

Esempio 6
P.M. Dott. ASTORI – **Allora signora lei pensa al marito, suo marito pensa a lei, però ognuno non pensa alla vita. Ma era così tremendo questo fastidio, così tremendo, eh? Mamma mia per 4 rumori**…
AVV. TROIANO – Io non vorrei che… Se lei potesse, allora cioè l'interrogatorio…
P.M. Dott. FADDA – Guardi, il tempo di caricare… Naturalmente diamo atto, visto che ci portiamo avanti e stiamo registrando, **diamo atto che i fatti che le si contestano sono gli stessi che lei ha sentito la volta scorsa**…
Indagata BAZZI ROSA ANGELA – Sì.
P.M. Dott. FADDA – Che sono nel verbale, gli elementi di prova sono quelli con una novità, che abbiamo detto anche a suo marito e che suo marito ha preso in considerazione, e cioè il fatto che proprio oggi ci hanno comunicato che quella macchia di sangue che c'era nella macchina, contiene non solo il sangue di Cherubini Valeria, ma misto anche il sangue di Frigerio Mario e quindi questo è un elemento…
P.M. Dott. ASTORI – Cominciamo, ha già letto tutto, **va bene allora**…
P.M. Dott. FADDA – **Quindi lei intende rispondere** con questo…
Indagata BAZZI ROSA ANGELA – Sì.
P.M. Dott. FADDA – **È avvertito che ha la facoltà di non rispondere**…
Indagata BAZZI ROSA ANGELA – Sì.
P.M. Dott. ASTORI – "Ho chiesto io di parlare con i magistrati perché voglio raccontare la verità", giusto?
Indagata BAZZI ROSA ANGELA – Giusto.
P.M. Dott. ASTORI – **Prego signora**, dica quello… **Incominci da dove vuole lei**.
[…]

Le battute riportate sono quelle che danno inizio all'interrogatorio. Il P.M. Astori, essendo il regista del potere interazionale, apre il colloquio e lo fa, sorprendentemente, creando una situazione di conflitto. Il suo primo intervento non può che essere percepito come provocazione. Questo confermerebbe quanto sostenuto da un filone di

ricerca per il quale in molti casi i motivi del conflitto non vanno ricercati al di fuori della situazione comunicativa, ma nelle scelte linguistiche e negli atteggiamenti che di volta in volta i partecipanti assumono (Orletti, 1994: 172).
La sequenza presenta poi una serie di battute in cui è possibile scorgere alcune delle fasi di quella *routine* conversazionale che opera nel soggetto una trasformazione di identità: la facoltà di non rispondere è data infatti solo agli indagati e agli imputati. L'individuo è costretto pertanto ad abbandonare la sua identità sociale per accogliere quella che gli viene attribuita dalla situazione istituzionale (Fele, 1997: 145).
Entrambe le situazioni viste mancano quindi di cortesia linguistica.
L'atto illocutivo finale del P.M. arriva inaspettatamente e ribalta la situazione sfavorevole dal momento che lascia all'indagata piena libertà di espressione. Ora, in tutti i tipi di interazione asimmetrica si può cogliere una specializzazione degli interagenti sulla base di un'associazione tra ruolo comunicativo e ruolo istituzionale. Pertanto, ad esempio, c'è chi può fare domande e chi invece deve limitarsi a rispondere (Orletti, 2000: 27). In tribunale, l'interrogato può contrapporre una domanda all'interrogante solo se deve chiedere spiegazioni preliminari necessarie per formulare una risposta adeguata (Bellucci, 2002: 158). Nel caso riportato Astori non pone domande, non si avvale della facoltà che gli spetta sebbene non rinunci al ruolo di regista, l'unico che ha il diritto di assegnare la parola. Questa strategia, al contrario della provocazione, rientra a pieno titolo nella prima fase di mediazione, sebbene solo in apparenza ridistribuisca equamente il potere all'interno dell'interazione (Biggio, 2005: 81).
Tutte le fasi d'intervento, ma principalmente la prima, richiedono al *counsellor* l'uso di un ascolto "attivo". Questi deve prestare la massima attenzione ai segnali che l'altro gli invia al fine di individuare eventuali problemi e, soprattutto, di rendere possibile una esplicitazione di elementi non espressi, o non espressi chiaramente, ma ritenuti fondamentali per l'analisi.

Esempio 7
[…]
P.M. Dott. ASTORI – Senta signora, delle cose che lei ha detto c'erano alcune cose vere ed alcune cose non vere. Ricominciamo bene, bene daccapo.
Indagata BAZZI ROSA ANGELA – Ok.
P.M. Dott. ASTORI – **Lei** innanzitutto **chi ha colpito per prima**?

Indagata BAZZI ROSA ANGELA – **Io**.
P.M. Dott. ASTORI – **Sì, ma chi ha colpito per primo**?
P.M. Dott. FADDA – **Quale persona**?
Indagata BAZZI ROSA ANGELA – **La Raffaella**.
[…]

Esempio 8
[…]
P.M. Dott. FADDA – Comunque ascolti lei **come li aveva chiamati i Carabinieri** quel giorno?
Indagata BAZZI ROSA ANGELA – Li avevamo chiamati **di notte** noi.
P.M. Dott. FADDA – **Sì, ma con il telefono di casa, con un cellulare**?
Indagata BAZZI ROSA ANGELA – **No, no il telefono di casa**, li avevamo chiamati ed i Carabinieri hanno detto, perché erano su che litigavano e tutto…
[…]

Tanto in *7* quanto in *8* l'indagata non comprende la domanda postale dai P.M. Alla sua risposta non pertinente tanto Astori quanto Fadda fanno seguire un *sì* come chiaro segnale di ricezione (Bazzanella, 2001[2]: 169), accompagnato però da un *ma* che riporta l'attenzione della signora al tema della domanda precedente: la persona che lei dice di aver colpito per primo nell'esempio 7, e con che cosa aveva chiamato i Carabinieri quella notte nell'esempio 8. In questo modo riescono a riportare l'interazione verso un principio di cooperazione conversazionale (Grice, 1978: 204); l'indagata avrebbe potuto anche avvalersi della facoltà di non rispondere. Nel procedimento penale solo il testimone non può avvalersi di tale facoltà.
La seconda fase applicativa di *counselling* è detta intermedia. È il momento in cui vengono riesaminati i problemi emersi, identificati debolezze e comportamenti inadeguati. Per far emergere le discrepanze in ciò che viene affermato e portare l'altro ad essere consapevole delle proprie contraddizioni, la guida mette in atto la strategia della confrontazione. Siamo in una fase molto delicata perché il cliente percepisce il *counsellor* ancora come un "intruso". Questi, per evitare un blocco relazionale, deve quindi attendere che gli vengano date microautorizzazioni prima di procedere con l'analisi:

Esempio 9
[…]
P.M. Dott. ASTORI – E il fumo **da dove** veniva?
Indagata BAZZI ROSA ANGELA - **Dall'appartamento** perché abbiamo lasciato la…, abbiamo aperto la porta ed è uscito tutto.

P.M. Dott. ASTORI – Sì, ma **il fumo quindi era già, era già forte**?
Indagata BAZZI ROSA ANGELA - **Sì, abbastanza**.
P.M. Dott. ASTORI – **Chi** ha fatto tutto il fuoco dentro, **lei o suo marito**?
Indagata BAZZI ROSA ANGELA – No, **il piumino con l'accendino** e ha cominciato a bruciare subito.
P.M. Dott. ASTORI – Va bene. Senta signora, **faccia uno sforzo, mi rendo conto che non è facile raccontare queste cose**, poi dopo le facciamo altre domande, ma spieghi bene adesso **in che punto è stato dato il fuoco e con che cosa** esattamente?
Indagata BAZZI ROSA ANGELA - Con niente.
P.M. Dott. ASTORI – **In che modo**, **come avete fatto**?
Indagata BAZZI ROSA ANGELA - **Con niente**. È solamente che **abbiamo preso il piumino** e lo…, cioè giornali, libri e il maglione della Raffaella e tutto quello che c'era nel cassetto, e **ho acceso con**…
P.M. Dott. ASTORI – Quale cassetto?
P.M. Dott. FADDA – **Se lei non fuma**.
Indagata BAZZI ROSA ANGELA – **L'accendino**.
P.M. Dott. FADDA – **Perché aveva l'accendino se non fuma**?
Indagata BAZZI ROSA ANGELA – **Me lo sono portata dietro.**
P.M. Dott. ASTORI – **Senta signora, 5 minuti fa lei aveva detto che era stato suo marito ad appiccare il fuoco**.
Indagata BAZZI ROSA ANGELA – **Anche lui l'avrà appiccat…, cioè può darsi che si è spento e…**
P.M. Dott. ASTORI – **No. Lei 5 minuti fa aveva detto che era venuto suo marito, ha visto e poi ha dato fuoco. No**?
Indagata BAZZI ROSA ANGELA – **Sì, sì**.
P.M. Dott. FADDA – Possiamo tornare indietro guardi.
Indagata BAZZI ROSA ANGELA – No, no è vero.
P.M. Dott. FADDA – **Signora, dovrebbe… Tanto a questo punto, una bugia non fa la differenza**.
P.M. Dott. ASTORI – **Chi ha dato fuoco**?
Indagata BAZZI ROSA ANGELA - Cioè **ho acceso**…
P.M. Dott. ASTORI – **Ho acceso, o ha acceso suo marito**?
Indagata BAZZI ROSA ANGELA - **No, ho acceso in cameretta e lui ha acceso in camera matrimoniale**.
[…]

Il principio di collaborazione che vige tra le parti nelle prime battute viene interrotto dalla domanda diretta del dott. Astori volta a stabilire chi ha appiccato il fuoco in casa della signora Castagna. La signora Bazzi cade in contraddizione e comincia ad opporre forti resistenze. Questa sua reazione poteva però essere prevista sulla base dei dati acquisiti. In questa fase del confronto, infatti, l'indagata addossa tutta la responsabilità dell'accaduto a sé, ribadendo più volte che il marito è estraneo ai fatti. Le figure guida mostrano palesemente di avere difficoltà nell'esercitare un controllo e un orientamento sulla

situazione. Solo tali difficoltà, d'altronde, potrebbero giustificare l'ultimo intervento del dott. Fadda. Per l'indagata «Una bugia non fa la differenza» poteva significare che l'accusa avesse già preso decisioni che non avrebbe potuto rimettere in discussione con eventuali nuove dichiarazioni. Per fortuna l'indagata non assume un atteggiamento di totale chiusura e alla fine chiama in causa il marito.
La natura del materiale di cui disponiamo al momento non ci consente di argomentare la fase finale dell'intervento di *counselling* in cui il cliente, assumendosi la piena responsabilità delle sue azioni, è chiamato in prima persona a valutare il proprio agire. In questo è aiutato dall'esperto che, adottando modalità comunicative appropriate, riesce ad allontanarlo dalla continua tensione tra libertà e carcerazione. Nel frammento seguente ci sembra però di poter cogliere un momento in cui la signora Bazzi si stava lasciando andare ad un processo di autoanalisi. Questo sempre ammettendo che non ci siano errori ascrivibili al trascrittore:

Esempio 10
[…]
P.M. Dott. ASTORI – Ed **il coltello dove l'ha preso**?
Indagata BAZZI ROSA ANGELA – **Quello da cucina**.
P.M. Dott. ASTORI – **Come è fatto**? Si ricorda **come**…
Indagata BAZZI ROSA ANGELA – **Cosa ho fatto**…
P.M. Dott. ASTORI – **Come** è fatto questo coltello, come era fatto?
Indagata BAZZI ROSA ANGELA – Normale, un coltello del ceppo.
P.M. Dott. ASTORI – Del ceppo?
Indagata BAZZI ROSA ANGELA – Sa quelli…
P.M. Dott. FADDA – Quindi nel ceppo ne manca uno.
Indagata BAZZI ROSA ANGELA – Sì, ma non di quelli lì che mancano dal ceppo, è un coltello sparso nella cucina così.
P.M. Dott. ASTORI – E li ha buttati via insieme ai vestiti?
Indagata BAZZI ROSA ANGELA – Sì.
[…]

All'interno dello scambio comunicativo il secondo turno della signora risulta affatto pertinente. L'assenza in Astori di un segnale di ricezione – presente negli esempi 7 e 8 – e di un segnale di interruzione rende il turno altamente significativo, perché del tutto inatteso, così come inattesa risulta a noi lettori la mancata svalutazione dell'informazione da parte del P.M. (De Beaugrande and Dressler, 1994: 157-79). La spiegazione di ciò potrebbe risiedere nei

puntini sospensivi che chiudono sia il turno di Astori, sia quello della signora Bazzi. Purtroppo, nelle trascrizioni giudiziarie i puntini rappresentano un segno d'interpunzione *passepartout* e quindi non rimandano ad una situazione specifica (Bellucci, 2002: 42)[5]. Nel nostro caso sembrerebbero segnalare una sovrapposizione di turni, piuttosto che un'interruzione. Pur volendo ammettere che il P.M. stia procedendo nell'interrogatorio secondo lo schema della sua agenda nascosta, vale a dire seguendo parametri e schemi che sfuggono a chi ha competenze diverse (Orletti, 2000: 29-30; Bellucci, 2002: 159-62), appare strano che si sia lasciato sfuggire un dato importante come questo. D'altronde, in ambito giudiziario si ha la compresenza di più agende nascoste e il porre domande serve anche a svelare l'agenda nascosta della controparte. In situazioni come queste abbiamo avvertito il limite di non aver potuto condurre un'analisi congiunta dei tre canali di comunicazione: verbale, non verbale e paraverbale.
Prima di concludere, richiamiamo l'attenzione sui comandi "nascosti" in negativo, che bisognerebbe sempre evitare se l'obiettivo è quello di ottenere una comunicazione efficace. Vediamo il perché:

Esempio 11
[…]
P.M. Dott. NALESSO – Il marito della Raffaella le ha fatto qualcosa di particolare?
P.M. Dott. FADDA – Dai signora lo dica, tanto…
Indagata BAZZI ROSA ANGELA – Niente, cioè mi spaventava.
P.M. Dott. FADDA – Cioè, come?
Indagata BAZZI ROSA ANGELA – Che voleva…
P.M. Dott. ASTORI – Vuole qualcosa da bere signora?
Indagata BAZZI ROSA ANGELA – **Io avevo paura**.
P.M. Dott. FADDA – Ma cosa le faceva per spaventarla? **Stia tranquilla** signora dai, **non si preoccupi**.
P.M. Dott. ASTORI – **Non si preoccupi** signora, noi siamo tutti un po' provati da questa situazione.
[…]

La difficoltà della signora a ritornare sulle dichiarazioni relative alle molestie subite dal signor Azouz Marzouk, porta i P.M. ad indirizzare all'indagata un'espressione rassicurante formulata come negazione – *non si preoccupi* – trascurando il fatto che le espressioni contenenti

[5] Per i problemi relativi alle trascrizioni rimandiamo a Bellucci and Palmerini (2004) e relativa bibliografia.

una negazione, e le negazioni stesse, appartengono alla struttura del linguaggio e non a quella del cervello. Per negare un'azione il cervello infatti deve prima processare l'evento in positivo, per poi provare a cancellarlo. Sottolineiamo 'provare' perché tale operazione risulta estremamente difficile (Knight, 2009: 206-07; Bandler, 2009: 289). Se invitiamo qualcuno a non pensare intensamente ad un elefante a *pois* rosa, questi se lo rappresenterà esattamente come gli abbiamo chiesto di non fare.

5 Conclusioni

Lo studio condotto ha messo in luce aspetti peculiari di un comportamento volto alla ritualizzazione del rapporto gerarchico esistente in società. Da un lato, i P.M. mettendo in primo piano la sfera sociale a cui appartengono non favoriscono uno scambio efficace tra i diversi stati dell'Io; dall'altro, gli indagati adattano la propria modalità espressiva a quella che di volta in volta viene loro richiesta, consapevoli di essere la parte debole di una comunicazione asimmetrica.

Ne consegue che nelle relazioni in cui il potere non è equidistribuito, o non è percepito come tale, emergono numerosi fattori che determinano scambi angolari o duplici che portano la parte forte a far ricorso a transazioni di tipo incrociato o angolare che sottendono significati nascosti differenti da quelli manifesti.

L'analisi di entrambi gli interrogatori ha evidenziato le difficoltà dei P.M nel gestire lo scambio comunicativo. Infatti, hanno spesso dovuto portare il proprio ruolo sull'Io Genitore, abbandonando quello Adulto; hanno perso il controllo della situazione in atto tanto da non percepire informazioni altamente rilevanti ai fini delle indagini; spesso mancavano di strumenti a cui far ricorso per superare tali difficoltà. L'Analisi Transazionale e il *Counselling* suggeriscono che una strategia comunicativa efficace avrebbe dovuto portare i P.M. a creare un clima di fiducia reciproca, per poi procedere ad una comunicazione tra stati dell'Io Adulto.

Le risultanze del nostro lavoro sembrano dunque confermare la necessità di un approccio multidisciplinare all'evento comunicativo, capace di offrire a chi lo gestisce una maggiore e più completa consapevolezza delle dinamiche interazionali che si instaurano tra gli interagenti. Di qui l'auspicio che si possano attivare in ambito

giudiziario percorsi formativi che sviluppino competenze necessarie a un uso efficace e consapevole della parola che è alla base della comprensione, dell'interpretazione e dell'applicazione della legge (Mariani Marini, 2003: IX).
De Mauro riflettendo sulla complessità e sulla natura del linguaggio umano scriveva

> […] val la pena che la linguistica sappia chiamare a occuparsi di linguaggio tutti coloro che sono interessati a esplorare scientificamente la storia e la natura degli esseri umani (De Mauro, 1998: 112).

Riferimenti bibliografici

Bazzanella, C. (2001[2][1994]) *Le facce del parlare. Un approccio pragmatico all'italiano parlato*. Firenze: La Nuova Italia.

Bazzanella, C. (2008) *Linguistica e pragmatica del linguaggio. Un'introduzione*. Roma-Bari: Laterza.

Bandler, R. (2009) *Il Potere dell'inconscio e della PNL*. Urgnano: NPL Italy – Alessio Roberti.

Bellucci, P. (2002) *A onor del vero. Fondamenti di linguistica giudiziaria*. Torino: UTET Libreria.

Bellucci, P. and M. Palmerini (2004) Processi penali: la prova si forma oralmente in Dibattimento, ma si deposita nella trascrizione degli atti, in *Quaderni del Dipartimento di Linguistica – Studi 14*, Università degli Studi di Firenze, Padova: Unipress, pp. 73-122.

Berne, E. (1971) *Analisi Transazionale e Psicoterapia*. *Un sistema di psichiatria sociale e individuale*. Roma: Astrolabio.

– (1979) *"Ciao!"... E poi?*. Milano: Bompiani.

Biggio, G. (2005) Counselling. *Metodi e applicazioni*. Roma: Gangemi.

Brown and Levinson (1987) *Politeness: Some Universals in Language Usage*. Cambridge: Cambridge University Press.

Cavicchioli, S. (1997) Processi in televisione, in P.P. Giglioli, S. Cavicchioli and G. Fele (eds) *Rituali di degradazione: Anatomia del processo Cusani*, Bologna: il Mulino, pp. 75-134.

De Beaugrande, R.A and W.U. Dressler (1994) *Introduzione alla linguistica testuale*. Bologna: il Mulino.

De Mauro, T. (1998) *Linguistica elementare*. Roma-Bari: Laterza.

Di Fabio, A. (1999) *Counseling. Dalla teoria all'applicazione*. Firenze: Giunti.

Fele, G. (1997) Strategie discorsive e forme della degradazione pubblica in tribunale, in P.P. Giglioli, S. Cavicchioli and G. Fele (eds) *Rituali di degradazione. Anatomia del Processo Cusani*, Bologna: il Mulino, pp. 135-227.

Grice, P. (1978) Logica e conversazione, in M. Sbisà (ed.) *Gli atti linguistici. Aspetti e problemi di filosofia del linguaggio*, Milano: Feltrinelli, pp. 199-219.

Guida, M.A. (1993) La maturità emotiva come fondamento di una corretta gestione del proprio ruolo, in A. Quadrio and D. Pajardi (eds) *Interazione e comunicazione nel lavoro giudiziario*. *Prospettive di Psicologia Giuridica* 3, Milano: Giuffrè, pp. 89-95.

Harris, T.A. (2000) *Io sono ok, tu sei ok*. Milano: Rizzoli.

Knight, S. (2009) *PNL al lavoro*. Urgnano: NLP Italy – Alessio Roberti.

Mariani Marini, A. (2003) Presentazione a *La lingua, la legge, la professione forense*, Consiglio Nazionale Forense, 2, Milano: Giuffrè, pp. IX-XXI.

Miglionico, A. (2000) *Manuale di Comunicazione e Couselling*. Torino: Centro Scientifico.

Moiso, C. and M. Novellino (1982) *Stati dell'Io. Le basi teoriche dell'Analisi Transazionale Integrata*. Roma: Astrolabio.

Orletti, F. (1994) Sulla superficie del conflitto, in F. Orletti (ed.) *Fra conversazione e discorso. L'analisi dell'interazione verbale*, Roma: La Nuova Italia Scientifica, pp. 171-84.

– (2000) *La conversazione diseguale*. Roma: Carocci.

Petrillo, G. (1994) Relazioni asimmetriche e strategie di influenza sociale: analisi di conversazioni medico-paziente, in F. Orletti (ed.) *Fra conversazione e discorso. L'analisi dell'interazione verbale,* Roma: La Nuova Italia Scientifica, pp. 99-119.

Quadrio, A. (1993) 'Introduzione a Interazione e comunicazione nel lavoro giudiziario'. *Prospettive di Psicologia Giuridica* 3. Milano: Giuffrè, pp. 3-13.

Vinella, P. and N. Lopez (2000) Il canale emotivo, in A. Miglionico (ed.) *Manuale di Comunicazione e Couselling*, Torino: Centro Scientifico, pp. 83-102.

Towards Automatic Dialogue Understanding

*Rodolfo Delmonte^, Antonella Bristot^, Vincenzo Pallotta**
^Department of Language Science, Università "Ca Foscari", 30123 Venezia,
*Department of Computer Science, Webster University, Geneva, Switzerland
delmont@unive.it, pallotta@webster.ch

Abstract

In this paper we will present work carried out to scale up the system for text understanding called GETARUNS, and port it to be used in dialogue understanding. The current goal is that of extracting automatically argumentative information in order to build argumentative structure. The long term goal is using argumentative structure to produce automatic summarization of spoken dialogues.
Very much like other deep linguistic processing systems (see Allen et al*, 2007), our system is a generic text/dialogue understanding system that can be used in connection with an ontology – WordNet – and other similar repositories of commonsense knowledge. Word sense disambiguation takes place at the level of semantic interpretation and is represented in the Discourse Model. We will present the adjustments we made in order to cope with transcribed spoken dialogues like those produced in the ICSI Berkely project. The low level component is organized according to LFG theory; at this level, the system does pronominal binding, quantifier raising and temporal interpretation. The high level component is where the Discourse Model is created from the Logical Form. For longer sentences the system switches from the top-down to the bottom-up system. In case of failure it will back off to the partial system which produces a very lean and shallow semantics with no inference rules.*
In a final section, we present preliminary evaluation of the system on two tasks: the task of automatic argumentative labelling and another frequently addressed task: referential vs. non-referential pronominal detection. Results obtained fair much higher than those reported in similar experiments with machine learning approaches.

1 Introduction

The system presented here has been achieved in over two decades with the goal of developing a broad-coverage, domain general natural language unterstanding system. The underlying grammar, lexicon, the semantics and all intermediate modules are intended to be domain-general and to be easily portable to different application domains. As

is the case with all rule-based systems, (but see also Allen *et al*, 2007), we have no need to collect and annotate corpora for specific subtasks because the system already has good performance in all current parsing and semantic related tasks (see Delmonte *et al*, 2006; Delmonte, 2007; 2008).
However, when we started last year to use the system to parse ICSI dialogues, we realized that the semantic representation and the output of the parser were both inadequate. So we worked at deficiencies we detected in an empirical manner. This approach made us aware of the peculiarities of spoken dialogue texts such as the ones made available in ICSI project (see Janin *et al*, 2003), and to the way to implement solutions in such a complex system. These dialogues are characterized by the need to argument in a exhaustive manner the topics to be debated which are the theme of each multiparty dialogue. The mean length of utterances/turns in each dialogue we parsed was rather long. This makes ICSI dialogues hard to compute. From a count of number of words x turn, we came up with the following average figures:

- percent of turns made of one single word: 30%;
- percent of turns made of up to three words: 40%;
- number of words x turn overall: 7;
- number of words x turn after subtracting short utterances: 11.

These values correspond to those found for PennTreebank corpus where we can count up to 94K sentences for 1M words – again 11 words per sentence. In analyzing ICSI, we found turns with as much as 54 words depending on the topic under discussion and on the people on the floor.
Computing semantic representations for spoken dialogues is a particularly hard task which – when compared to written text processing – requires the following additional information to be made available:

- adequate treatment of fragments;
- adequate treatment of short turns, in particular one/two-words turns;
- adequate treatment of first person singular and plural pronominal expressions;
- adequate treatment of disfluencies, thus including cases of turns made up of just such expressions, or cases when they are found inside the utterance;

- adequate treatment of overlaps;
- adequate treatment of speaker identity for pronominal coreference.

In addition, in our system every dialogue turn receives one polarity label, indicating negativity or positivity, and this is computed by looking into a dictionary of polarity items. This is subsequently used to decide on argumentative automatic classification.
We will address each such topics in what follows. The paper is organized with a presentation of the system in section 2; then problems will be presented and discussed by addressing their computational role and the module responsible for its treatment, in the following section. We will then present some data and an evaluation. A final section will be devoted to conclusions.

2 The System GETARUNS

GETARUNS[1], the system for text understanding developed at the University of Venice, is organized as a pipeline which includes two versions of the system: what we call the Partial and the Deep GETARUNS. At first we will present the Deep version, which is equipped with three main modules: a lower module for parsing, where sentence strategies are implemented; a middle module for semantic interpretation and discourse model construction which is cast into Situation Semantics; and a higher module where reasoning and generation takes place.
The system is based on LFG theoretical framework (see Bresnan, 2000) and has a highly interconnected modular structure. The Closed Domain version of the system is a top-down depth-first DCG-based parser written in Prolog Horn Clauses, which uses a strong deterministic policy by means of a lookahead mechanism with a WFST to help recovery when failure is unavoidable due to strong attachment ambiguity.
It is divided up into a pipeline of sequential but independent modules which realize the subdivision of a parsing scheme as proposed in LFG theory where a c-structure is built before the f-structure can be projected by unification into a DAG (Direct Acyclic Graph). In this sense we try to apply in a given sequence phrase-structure rules as

[1] The system has been tested in STEP competition (see Delmonte, 2008), and can be downloaded at, http://project.cgm.unive.it/html/sharedtask/.

they are ordered in the grammar: whenever a syntactic constituent is successfully built, it is checked for semantic consistency. In case the governing predicate expects obligatory arguments to be lexically realized they will be searched and checked for uniqueness and coherence as LFG grammaticality principles require.

Syntactic and semantic information is accessed and used as soon as possible: in particular, both categorial and subcategorization information attached to predicates in the lexicon is extracted as soon as the main predicate is processed, be it adjective, noun or verb, and is used to subsequently restrict the number of possible structures to be built. Adjuncts are computed by semantic compatibility tests on the basis of selectional restrictions of main predicates and adjuncts heads.

The output of grammatical modules is fed then onto the Binding Module which activates an algorithm for anaphoric binding. Antecedents for pronouns are ranked according to grammatical function, semantic role, inherent features and their position at f-structure. Eventually, this information is added into the original f-structure graph and then passed on to the Discourse Module (hence DM).

The grammar is equipped with a core lexicon containing most frequent 5,000 fully specified inflected word forms where each entry is followed by its lemma and a list of morphological features, organized in the form of attribute-value pairs. However, a morphological analyser for English is also available with big root dictionaries (25,000 for English) which only provide for syntactic subcategorization, though. In addition to that, there are all lexical form provided by a fully revised version of COMLEX, and in order to take into account phrasal and adverbial verbal compound forms, we also use lexical entries made available by UPenn and TAG encoding. Their grammatical verbal syntactic codes have then been adapted to our formalism and are used to generate a subcategorization scheme with an aspectual and semantic class associated to it – however no restrictions can reasonably be formulated on arguments of predicates. Semantic inherent features for Out of Vocabulary Words, be they nouns, verbs, adjectives or adverbs, are provided by a fully revised version of WordNet – plus EuroWordnet, with a number of additions coming from computer, economics, and advertising semantic fields – in which we used 75 semantic classes similar to those provided by CoreLex. The complete lexicon counts 270K entries.

When each sentence is parsed, tense aspect and temporal adjuncts are accessed to build the basic temporal interpretation to be used by the temporal reasoner. Eventually two important modules are fired: Quantifier Raising and Pronominal Binding. QR is computed on f-structure which is represented internally as a DAG. It may introduce a pair of functional components: an operator in the structure where the quantifier can be raised, and a pool containing the associated variable where the quantifier is actually placed in the f-structure representation. This information may then be used by the following higher system to inspect quantifier scope. Pronominal binding is carried out at first at sentence internal level. DAGs will be searched for binding domains and antecedents matched to the pronouns if any to produce a list of possible bindings. Best candidates will then be chosen.

2.1 The Upper Module

GETARUNS has a highly sophisticated linguistically based semantic module which is used to build up the DM. Semantic processing is strongly modularized and distributed amongst a number of different submodules which take care of Spatio-Temporal Reasoning, Discourse Level Anaphora Resolution and other subsidiary processes like Topic Hierarchy which cooperate to find the most probable antecedent of coreferring and cospecifying referential expressions when creating semantic individuals. These are then asserted in the DM, which is then the sole knowledge representation used to solve nominal coreference, before proceeding to access external knowledge in the ontologies. The system uses two resolution submodules which work in sequence: they constitute independent modules and allow no backtracking. The first one is fired whenever a free sentence external pronoun is spotted; the second one takes the results of the first submodule and checks for nominal anaphora. They have access to all data structures contemporarily and pass the resolved pair, anaphor-antecedent to the following modules. Semantic Mapping is performed in two steps: at first a Logical Form is produced which is a structural mapping from DAGs onto unscoped well-formed formulas. These are then turned into situational semantics informational units, infons which may become facts or "sits". Each unit has a relation, a list of arguments which in our case receive their semantic roles from lower processing – a polarity, a temporal and a spatial location index. The

clause-level interpretation procedure interprets clauses on the basis of lexical properties of the governing verb. This is often non available in short turns and in fragments. So in many cases, fragments are built into a sentence by inserting a dummy verb which varies from dummy BE or dummy SAY depending on speech act present.

3 The Spoken Dialogue Additions

We will proceed by addressing each problem presented above in the order with which it is coped with by the system.

3.1 The Algorithm for Overlaps

Overlaps are an important component of all spoken dialogue analysis (Delmonte, 2003). In all dialogue transcriptions, overlaps are treated as a separate turn from the one in which they occur, which usually follows it. This is clearly wrong from a computational point of view. For this reason, when computing overlaps we set as our first goal that of recovering the temporal order in which speaker and interlocutor interact. This is done because:

- overlaps may introduce linguistic elements which influence the local context;
- eventually, they may determine the interpretation of the current utterance.

For these reasons, they cannot be moved to a separate turn because they must be semantically interpreted where they temporally belong. In addition, overlaps are very frequent. The algorithm we built looks at time stamps, and every time the following turn begins at a time preceding the ending time of current turn it enters a special recursive procedure. It looks for internal interruption in the current turn and splits the utterance where the interruption occurs. Then it parses the split initial portion of current utterance and continues with the overlapping turn. This may be reiterated in case another overlap follows which again begins before the end of current utterance. Eventually, it returns to the analysis of the current turn with the remaining portion of current utterance.
In Tab. 1 below we present data related to overlaps for the first 10 dialogues we computed. We classified overlaps into two types – WHILE and AFTER – according to whether they take place inside the

turn of the current speaker or at the end, the second case being regarded as normal and non disrupting of the current speaker's conversational plan.

	total	Continue	Interr-upt	inter_cont	inter_ change	inter_ other
turns	**13158**	-	-	-	-	-
while	**1624**	**1369**	**46**	**87**	**22**	63
after	1461	-	-	-	-	-

Table 1: Overlaps and their effects on Planning

On a total number of 13,158 turns we thus computed 3,085 overlaps divided up nicely almost half and half for each of the two classes. Then we proceeded by subdividing WHILE overlaps into 5 subclasses where Continue indicates the current speaker continues talking; Interrupt, the current speaker is interrupted and there is no continuation; Inter_Cont, the current speaker is interrupted but then Continues his/her plan in a following turn; Inter_Change, the current speaker is interrupted and changes his/her plan, but either changing subject topic, or answering the overlapper. Eventually we had Inter_Other, which indicates cases in which dialogue is interrupted by other speakers.

As can be easily noticed, the case constituted by Inter_Change, which is the most interesting from a semantic and pragmatic point of view is in fact the less frequent. We assume, however, that this may be determined by other factors attaining to the type of conversation being entertained by the participants, as well as by the nature of the topics discussed, and eventually by the personalities of the interlocutors.

3.2 The Treatment of Fragments and Short Turns

Fragments and short turns are filtered by a lexical lookup procedure that searches for specific linguistic elements which are part of a list of backchannels, acknowledgements expressions and other similar speech acts. In case this procedure has success, no further computation takes place. However, this only applies to utterances shorter than 5 words, and should be made up only of such special words. No other linguistic element should be present apart from non-words, that is

words which are only partially produced and have been transcribed with a dash at the end:

- graceful failure procedures for ungrammatical sentences, which might be fullfledged utterances but semantically uninterpretable due to the presence of repetitions, false starts and similar disfluency phenomena. Or else they may be just fragments, i.e. partial or incomplete utterances, hence non-interpretable as such; this is done by imposing grammatical constraints of wellformedness in the parser;
- failure procedures for utterances which are constituted just by disfluency items and no linguistically interpretable words. These must be treated as semantically empty utterances and are recognizable by the presence of orthographic signs indicating that the word/s have not been completed and are just incomprehensible; this is done by inspecting the input in search of special orthographic marks and preventing the utterance to be passed down to the partial/deep parser.

On the contrary, we implemented a principled treatment of elliptical utterances which contribute one specific speech act or communicative act. They may express agreement/disagreement, acknowledgements, assessments, continuers etc. All these items are computed as being complements of abstract verb SAY which is introduced in the analysis, and has as subject, the name of current speaker.

4 Automatic Argumentative Annotation

At first we shall provide a state of the art and then we shall comment in detail our approach.

4.1 Detecting Argumentative structure – issues and theories

As shown by Rosemberg and Silince (1999), tracking argumentative information from meeting discussions is of central importance for building summaries of project memories since, in addition to the "strictly factual, technical information", these memories must also store relevant information about decision-making processes. In a business context, the information derived from meetings is useful for future business processes, as it can explain phenomena and past decisions and can support future actions by mining and assessment (Pallotta *et al*, 2004).

In a section below we will describe in detail how the annotation process takes place. Here we want to highlight the main features of

this process. This first level of annotation is based on the shallow dialogue model, proposed in Armstrong (2003), of which it is a modified version. This model provides a simple operational structure of dialogues based on three categories:

• ***a dialog*** is a non empty set of episodes; a new episode is identified by a topic/speaker shift;
• ***an episode*** is a non empty set of turns; turns are individuated at prosodic level – more on turns below;
• ***a turn*** is a non empty sequence of clauses/utterances and their boundary is a long pause.

In addition to the shallow dialogue model, we consider the adoption of a deeper structured representation based on argumentation theory. We assume that meeting dialogues are better viewed from the Collaborative Decision Making (CDM) perspective. In CDM, a meeting is defined as a multi-party (multi-agent) decision making process: a collaborative process, where agents follow a series of communicative actions in order to establish a common ground on the dimension of the problem. The main four dimensions of CDM process are:

• an overall task issue;
• a set of alternative proposals;
• a set of arguments in favour or against each proposals;
• a collection of choice criteria (perspectives and preferences) settled upon the participants;
• a decision (or evaluation) function that combines criteria to judge the alternatives.

This definition focuses on the processes, which take place during meetings and how these processes contribute to the accomplishment of a joint goal. In order to capture the above dimensions, we then adopted and extended a suitable argumentative model of discussions, namely the IBIS model proposed by (Kunz and Rittel, 1970). The IBIS model provides us with an abstract description of the discussion's rationale by outlining the important points discussed, the conflicts arisen and, hopefully solved, and the decisions that have been made. The IBIS model abstracts from the dynamics of the discussion, which needs to be modelled as well in order to extract the

IBIS structures from meeting events. Relevant meeting events are special types of Dialogue Acts that have an argumentative force. This type of Dialogue Acts (Bunt, 1979), called Argumentative Acts, are backward-looking acts with forward-looking expectations (Goffman, 1981).

Within the Adjacency Pairs model (Schegloff and Sacks, 1973), the importance of tracking agreement and disagreement in discussions has been recognized also in (Galley *et al*, 2003; Hillard, Ostendorf, and Shriberg, 2003). Although these methods have the great advantage of being automatic, they only partially help in reconstructing the argumentative information we need in order to answer real user queries. This model has been adopted by (Niekrasz *et al*, 2005) for the real-time reconstruction of an argumentative structure by overhearing discussions in design meetings. Finally, Rienks and Verbree (2006) propose the Twente Annotation Schema, that is based on fewer categories but more relation types being inspired by the Rhetorical Structure Theory (Mann and Thompson, 1988).

The argumentative structure defines the different patterns of argumentation used by participants in the dialogue, as well as their organization and synchronization in the discussion. The limits of sequential analysis of conversation (Schegloff and Sacks, 1973) have been already pointed out by Goffman (1981), who proposes to extend the notion of adjacency pair with that of chains of interaction rounds. As for other related work, we also see similarities of our approach with the argumentation dependency grammar proposed by Lo Cascio (1991), although in his work only argumentative structure of monologues is considered. In fact, when analyzing dialogues, adjacency pairs are not enough to represent the hierarchical structure of the discussion: consider, for instance, an answer that replies to two different questions in the discussion. In this case, we need to add a relation that links the answer to both of the questions. We call this relation "replies_to". The "replies_to" links a (re)action to one or more previous (possibly in time) actions and induces an argumentative chain structure on the dialogue, which is local to each action and which enables the visualization of its context. For instance, the context of the action of "accepting a clarification" will be a chain of linked actions, namely the action of the clarification, that of the proposal that is clarified and the action of raising an issue for which the proposal was made. Argumentative actions can overlap in time, as, for instance,

in those cases where the acceptance of a justification is uttered in the form of "backchannel" during the presentation of the justification.
Argumentative actions such as REQUEST, ACCEPT, REJECT might correspond to basic dialogue acts (Clark and Popescu-Belis, 2004). In this case we have refined the concept of dialogue act and adjacency pairs by specifying the role of dialogue acts in constructing the argumentative structure of the discussion through the "replies_to" relation.
When using the IBIS mark-up labels, a meeting is decomposed into several stages such as issues, proposals, and positions, each stage being possibly related to specific aggregations of elementary dialogue acts. Moreover, argumentative interactions may be viewed as specific parts of the discussion where several dialogue acts are combined to build such an interaction; as, for instance, a disagreement could be seen as an aggregation of several acts of reject and accept of the same proposal. From this perspective, we elaborated an argumentative coding scheme, the Meeting Description Schema (Pallotta *et al*, 2004), which takes into account the different stages (or episodes) defined by the IBIS model and extend the concept of adjacency pairs to relate these episodes to each other and to the corresponding argumentative function.
In MDS, the argumentative structure of a meeting is composed of a set of topic discussion episodes (a discussion about a specific topic). In each discussing topic, there exists a set of issue discussion episodes. An issue is generally a local problem in a larger topic to be discussed and solved. Participants propose alternatives, solutions, opinions, ideas, etc. in order to achieve a satisfactory decision. Meanwhile, participants either express their positions and standpoints through acts of accepting or rejecting proposals, or by asking questions related to the current proposals. Hence, for each issue, there is a corresponding set of proposals episodes (solutions, alternatives, ideas, etc.) that are linked to a certain number of related positions episodes (for example, a rejection to a proposed alternative in a discussing issue) or questions and answers.

4.2 Our Approach

Automatic Argumentative Annotation is carried out by a special module activated at the very end of the computation of the each dialogue. This module takes as input the complete semantic representation produced by the system recorded in Prolog facts in the DM. The elements of semantic representation we use are the following ones:

- all facts in Situation Semantics contained in the Discourse Model, which include individuals, sets, classes, cardinality, properties related to entities by means of their semantic indices;
- facts related to spatiotemporal locations of events with logical operators and semantic indices;
- vectors of informational structure containing semantic information at propositional level, computed for each clause;
- vectors of discourse structure with discourse relations computed for each clause from informational structure and previous discourse state (for an evaluation of system's performance, see Delmonte, Nicolae and Harabagiu, 2007);
- dialogue acts labels associated to each utterance or turn following ICSI classification;
- overlaps information computed at utterance level;
- topic labels associated to semantic indices of each entity marked as topic of discourse;
- all utterances with their indices as they have been automatically split by the system.

To produce Argumentative annotation, the system uses the following 21 Discourse Relations labels:

statement, narration, adverse, result, cause, motivation, explanation, question, hypothesis, elaboration, permission, inception, circumstance, obligation, evaluation, agreement, contrast, evidence, hypoth, setting, prohibition.

These are then mapped onto five general argumentative labels. In addition we use the label DISFLUENCY for all those turns that contain fragments which are non-sentences and are semantically uninterpretable.

ACCEPT, REJECT/DISAGREE, PROPOSE/SUGGEST, EXPLAIN/JUSTIFY, REQUEST, DISFLUENCY.

The algorithm works in the following manner:

1. It recovers Dialogue Acts for each dialogue turn as they have been assigned by the system. These labels coincide with ICSI labels (BKC, ACK, FGB, FHD, RHQ – that is Floor Grabber, Floor Holder, Backchannel, Acknowledge, RhetoricQuestion – with the addition of NEGation, ASSent, MTVation, PRPosal, GRTeeing, CNLusion;
2. It recovers Overlaps as they have been marked during the analysis;
3. It produces an Opinion label which we call Polarity, which can take one of two values: Positive or Negative according to whether the sentence contains positive or negative linguistic descriptions;
4. It produces a list of Hot Spots and builds up Episodes, where Hot Spots is simply a set of turns in sequence where the interlocutors overlap each other frequently. Episodes on the contrary are a set of turns in which a single speaker "arguments" his/her topics which may occasionally be interrupted by overlaps or by short continuers, backchannel or other similar phenomena by other speakers without however grabbing the floor;
5. Then the main predicate that assigns argumentative labels is called:
   i. at first it tries exceptions on the basis of the actual words contained in the turn. These exceptions may be constituted by Greetings, specific Speech Acts, Conventional utterances pronounced in specific situations like Thanking, etc.;
  ii. then Short utterances are checked. In case they end up with a question mark they are labeled as Questions. Else, the Dialogue Act label is considered. Negations are also computed here;
  iii. now the main call is activated. In order to start matching the rules, the semantic information is recovered for the current turn, clause by clause;
   iv. when semantic information has been recovered the rules are fired. There are some 33 rules which take as input the following vector of features:

assignargument(NoCl, [Pol,DialAct], DiscDom, DiscRel, Relev, DomPointView, Output)

where Output is the output label chosen by the rule; DiscDom may be Factive or NonFactive, Suggestion or Proposal; Relevance may be foreground or background; DomPointView may be objective or subjective. Rules are applied by matching input labels in a FSA manner. However, sometimes conditions and constraints are made to apply. For instance, *analyzecontext(NoCl)*, checks to verify whether the current speaker holds the floor in the 2 preceding or following clauses.

   v. the rules produce a set of argumentative labels, one for each clause. The system then chooses the label to associate to the turn utterance from a hierarchy of argumentative labels graded for Pragmatic Relevance which

establishes that, for instance, Question is more relevant than Negation, which is more relevant than Raise Issue, etc.

A short example extracted from Dialogue 1 is reported in the Appendix at the end of the paper where we report for each utterance the Predicate-Argument structure and the governing predicate/s preceded by the argumentative label chosen by the system. Eventually, we are able to evaluate the degree of collaboration vs. competitiveness of each participant in the conversation and make a general statement like this one produced automatically for Dialogue 1,

> The speaker that has held the majority of turns is
> - Don with a total of 512 turns,
> followed by
> - Morgan with a total of 456;
> The speaker that has undergone the majority of overlaps is Morgan followed by Jane;
> The speaker that has done the majority of overlaps is Morgan followed by Jane;
> Morgan is the participant that has been most competitive.

The system has been used to parse the first 10 dialogues of the ICSI corpus for a total number of 98,523 words and 13,803 turns. This has been done to "train" the system: what happened was that, for the first 5 dialogues, we had to take care of failures. We also had to tune all the modules and procedures carefully. In particular, the module for argumentative automatic classification was incrementally improved in order to cover all conventional ways to express Agreement. For this reason, we then chose two random additional dialogues to test this second task.

4.3 Experimental Results

We had one skilled linguist to provide a turn level annotation for argumentative labels: we don't have any agreement measure in this case, even though we expect the annotation to be in line with current experiments on the same subject (Pallotta, Seretan and Ailomaa, 2007). In the following table we report data related to the experiment of automatic annotation of argumentative categories. On a total of 2,304 turns, 2,251 have received an argumentative automatic classification, with a Recall of 97.53%. As can be gathered from Tab.

2., the F-score is fairly high compared to current results reported in the literature on the same topic which are all below 80%.

	Correct	Incorrect	Total Found
Accept	662	16	678
Reject	64	18	82
Propose	321	74	395
Request	180	1	181
Explain	580	312	892
Disfluency	19		19
Total	1826	421	2247

Table 2: Overall count of argumentative labels

We computed Precision as the ratio between Correct Argumentative Labels/Found Argumentative Labels, which corresponds to 81.26%. The F-score is 88.65%.

5 The Anaphora Resolution Module

The problem represented by pronominal expressions in dialogues needs to be addressed fully and not by means of ad hoc solutions. This requires a full-fledged system for anaphora resolution. One such system is shown in Fig. 1, where we highlight the architecture and main processes undergoing at the anaphora level. First of all, the subdivision of the system into two levels: Clause level – intrasentential pronominal phenomena – where all pronominal expressions contained in modifiers, adjuncts or complement clauses receive their antecedent locally. Possessive pronouns, pronouns contained in relative clauses and complement clauses choose preferentially their antecedents from list of higher level referring expressions. Not so for those pronouns contained in matrix clauses. In particular, the ones in subject position are to be coreferred in the discourse. This requires the system to be equipped with a History List of all referring expressions to be used when needed. In the system, three levels are indicated: Clause level, i.e. simple sentences; Utterance level, i.e. complex sentences; Discourse level, i.e. intersententially.

Our system computes semantic structures in a sentence by sentence fashion and any information useful to carry out anaphoric processes

needs to be made available to the following stretch of dialogue.

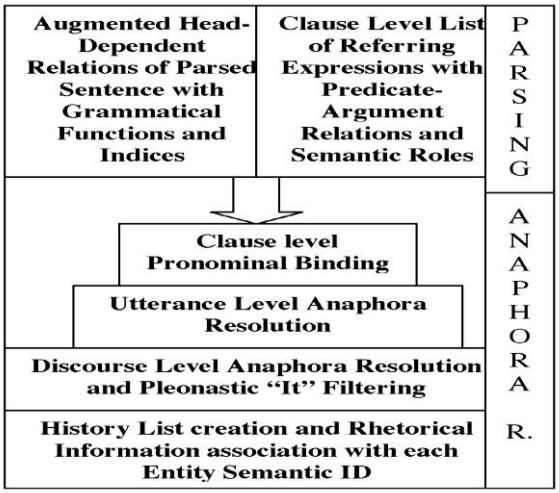


Figure 1: Anaphoric Processes in GETARUNS

6. The Experiments

We set up a number of experiments in order to test the new version of the system on the task of detecting referential from non-referential uses of personal pronouns YOU, WE and the pronoun IT.

6.1 State of the Art

Although much has been written about the referentiality/non-referentiality of pronouns in written text (mostly the third person neutral pronoun *it*), only recently the interest has shifted on spoken dialogues, partly thanks to the availability of annotated corpora. The main problems when trying to work with spontaneous speech are constituted by the presence of disfluencies, hesitations, abandoned utterances, interruptions, backchannels, etc.
According to Strube and Müller (2003), another problem is represented by the fact that spontaneous speech contains more instances of referential and non-referential pronouns than written text and also that spontaneous speech is characterized by the presence of a large number of singletons among the pronouns that are usually expletive or *vague* and cause problems for pronoun resolution algorithms which usually attempt to always find an antecedent.

6.1.1 You

Very few studies deal with the analysis on the referential/non-referential use of *you* in spoken dialogue. Y*ou* is an ambiguous pronoun because it can either be generic, i.e. does not refer to an addressee or to one of the participants in the dialogue, or referential and in this case, if singular, refers to the addressee of current utterance; else in case it is plural, it has more than one antecedent in previous discourse stretch. In their experiments for the resolution of *you* based on a portion of the Switchboard Corpus consisting of two-party telephone conversations, Gupta, Purver and Jurafsky (2007) distinguish between the non-referential and referential uses of *you* and also provide a few hints at the identification and classification of some idiomatic expressions, such as *you know* or clusters like *two of you*, *some of you*, etc., where *you* is always generic. Their analysis, primarily based on two-person conversations, has recently been applied quite successfully to multi-party dialogues.

According to the authors, in multi-party dialogues it is very important to identify the speaker and determine whether he is directly referred to using a second person pronoun or not. In spoken dialogue, in fact, most cases of *you* seem to be generic. They identify three types of *you*: generic, referential and reported referential (when *you* is mentioned in reported speech). They extract a number of features (Sentential, POS, Dialogue Act and Oracle Context features) directly derived from the corpus; the context features carry information (correct generic/referential label) about the uses of *you* in preceding context. Also some other features that may be of importance in detecting the correct role of *you* in dialogues, i.e. the presence of a question mark, are taken into consideration.

Even though the overall performance of the system is quite good with high precision values, 94% for the highest-accuracy result with only 36 errors out of 673 cases found, and accuracy by using all sets of features reaching a value of 84.4%, the authors realize how the use of some features, especially the context feature (accuracy=45.66%), are non influent in the analysis, while others perform quite well (dialogue act=80.92%). This is motivated by the fact that the generic/referential status of *you* may be independent of previous *yous* and that only looking at the neighbouring context may not be sufficient for a correct

analysis of this pronoun. They also checked some prosodic features which apparently did not add useful information.
In a second analysis performed on a 15-meeting subset of the AMI Meeting Corpus, the classification of *you* resulting from the annotation of the spoken text is of four kinds: generic, referential, reported speech and discourse marker (used to mark cases of *you* in idiomatic expressions and clusters as mentioned above). The authors decided to focus mainly on the categories generic/referential because the others occur in less than 2% of the dataset. For multi-party dialogues the result are worse than the two-speaker ones, resulting in a value of accuracy around 75% when using all the features derived from their previous work. They realize how multi-party dialogues are much more complex than two-party dialogue. A further analysis is then made only on the cases of referential *you* with the aim of identifying the addressee; this is achieved through a four-value model where value 1 is given to the potential addressee and value 4 addressing the entire group of speakers. The analysis is based on three sets of features (structural, durational and lexical) and the system reaches an accuracy of 47.2%.
Javanovich and op den Akker (2004) in their study for the addressee detection in face to face meetings of the MRDA Corpus, notice how *you* may often refer to meeting participants excluding the speaker of current utterance. The usage of quantifying determiners, numerals and infinitive pronouns may help in distinguishing *you* as a specific person from *you* as a group. At the same time, first name detection can be a very important means for addressee determination, especially when used in the vocative form. They also noticed that the speaker of utterances marked with question tags usually speaks directly to the addressee to provide information. This may be very important whenever there is a need to identify the addressee.
The method illustrated by Javanovich and op den Akker (2004) seems to be quite good in determining addressee identification reaching an accuracy of 83.74% also because they use information derived from the study of some important non linguistic features such as gaze direction, gesture and context.

6.1.2 We

To our knowledge studies on the referential/non referential nature of the pronoun *we* in spoken text are rare. Javanovich and op den Akker (2004) merely mention *we* observing how in multi-party dialogues *we* can either be used to refer to a subgroup of people taking part to the meeting or to all the participants, rather than to a single person. This is quite evident in open-ended questions that do not require a specific answer and which are usually addressed to all meeting participants; while this is not true for *you*, whenever an open question contains this pronoun, it is usually referred to a single speaker.

6.1.2 It

Several studies deal with the automatic classification of this pronoun, even though only a very few of them focus on the role of *it* in spoken dialogues. The problem with the pronoun *it* is two-fold: *it* can be generic, referential and expletive; moreover, *it* can have NP- and non-NP-antecedents, especially in spoken dialogues.
In their study Strube and Müller (2003) show how *it* can be classified as referential or non-referential by taking into consideration the local context surrounding the pronoun. Spoken text contains more pronouns with non-NP-antecedents than written text; specifically, *it* is a *vague* pronoun that may refer to different kinds of abstract objects from previous discourse: sentential antecedents or VP-antecedents. In their work, this class of vague pronouns also includes expletive pronouns, which are non-referential at all, and the so-called *vague anaphors*, for which humans don't seem to be able to determine an antecedent.
Their analysis, operated by a system originally used on written text and extended with a set of features designed for spoken dialogue, is based on part of the Switchboard Corpus that consists of 16,601 markables (sequence of words and attributes associated with them). Most of the pronouns found in the corpus are singletons which are vague or expletive and cause major problems in the processing of reference resolution algorithms. Strube and Müller's recognition is based on two classes of features used for written textxs: NP-level (grammar level) and Coreference-level features to which a new group of feature for the analysis of spoken dialogues has been added which include the type of antecedent (NP, VP, S) or the preference that some

verbs exhibit for arguments of different types. The study is conducted on three sets of third person pronouns: feminine and masculine, plural and neuter with a special interest in the latter since the neuter pronouns may have non-NP-antecedents. The results are quite good for the feminine, masculine and plural pronouns, but quite low for the neuter pronoun (~40.00% precision/~8.72% recall/~14.68% f-measure) and this is mostly due to the presence of many singletons and vague pronouns which do not have an antecedent and for which their model finds one nonetheless. The values of the analysis on a total of 1250 pronouns are: precision=56.74, recall=40.72, f-measure=47.82.

In a more recent work, Müller (2006) presents an implemented version of the system which was developed on the basis of meeting dialogues of the ICSI corpus. The task described is to identify and filter out non-referential cases of *it*; in this new version of the system there's no a priori knowledge available as to whether *it* is referential or not. According to the author, the best approach for the recognition of non-referential *it* should be based on filtering conditions and not solely on selection conditions.

The instances of *it* in five of the dialogues of the ICSI Corpus have been manually annotated by two human annotators and classified as belonging to six different classes: *normal*, *vague* – two sub-types of referential *it* – and *discarded*, *extrapos it*, *prop-it* and *other* – used to define cases of non-referential *it*. After this annotation, *K scores* have been calculated in order to check the reliability of the annotation. On a total of 1,017 examples of *it*, 62.5% are referential. After some editing and preprocessing, the dialogues have been classified according to a vector of features: some represent the syntactic patterns of the text capturing the left and right context of *it*, others contain lexical information about the predicative content of *it*, a third group of features captures a wider context of *it*, finally a fourth group contains some features used in order to check whether the pronoun is preceded by a preposition or if *it* appears as subject of some verbs like *seem*, *look*, *appear*, etc. In a second step machine learning techniques have been applied to detect non-referential *it*. The good results obtained by using information about interrupting points and sentence boundaries (precision:80.0%, recall:60.9%, f-score:69.2) is, according to the author, a sign of the utility of classifying *it* as non-referential in multi-party dialogues. Also the lower results obtained by simply employing

automatically determined features show how the use of filtering procedures for non-referential *it* represents an easy way to deal with the problem.
In their recent work, Bergsma *et al* (2008: 4) focus on *it* because, as they point out «*it* is one of the most frequent words in the English language, accounting for about 1% of tokens in text and over a quarter of all third-person pronouns[2]. Usually between a quarter and a half of *it* instances are non referential». Although not tested on spoken text, they used the Google Web 1T 5-gram Corpus, they present an analysis that may provide some good hints at the classification of *it* in spoken language. Their system, COMBO, is based on methods that do not rely heavily on hand-crafting of specific features like many other systems tend to do. COMBO deals mainly with the extraction of *context patterns* where some *pattern fillers* (i.e. words that can take the place of the pronoun) are extracted from a large number of n-grams (to a maximum of 5-grams) in order to determine which *it* can be replaced by an NP and which cannot; to identify non-referential *it* they count how often the pronoun appears as a pattern filler. They show how the position of *it* in the context pattern is usually fundamental in determining the referential/non-referential nature of *it*. Compared to other systems, COMBO performs quite well as for F-score values. Unfortunately, the system seems to detect only pronouns that are referential to a sentence and not to a noun – hence the necessity to extend the analysis to preceding discourse – and, above all, if the position of the pattern filler cannot be occupied for various reasons a possible referential occurrence of the pronoun is lost.

7 The Experiment: Method and Discussion

This work has tried to solve problems related to referential vs generic interpretation of pronominal expressions "it", "you" and "we".
In order to take decisions as to whether pronouns are to be interpreted as referential or not, a recursive procedure checks the type of governing predicate. Referential pronouns are then passed on to the pronominal binding algorithm that looks for local antecedents if any. Otherwise, the pronouns is labelled as having External coreference in the previous discourse stretch. The Anaphora Resolution module will

[2] *E.g*. http://ucrel.lancs.ac.uk/bncfreq/flists.html.

then take care of the antecedent and a suitable semantic identifier will be associated to it. On the contrary, if the pronouns are judged to be referentially empty or generic, no binding takes place. The pronoun is associated to the label "arbitrary" which prevents the system from computing it as a referential expression in the following semantic interpretation module. Not being a semantically interpretable linguistic element is important in order to avoid the pronoun from entering the Topic Hierarchy Centering module and being assigned a score. At the same time, the pronominal linguistic expression has to be used in order to complete predicate-argument structure. We assume that using such referentially empty pronouns corresponds to using expletive "it" in agentless passive constructions, or expletive "there" in presentative sentences.

To satisfy these constraints we let the parser compute these generic pronouns as PREDs and be associated to an f-structure. Then the pronominal binding module associates the label "arbitrary" to the value INTERPRETATION (it can be quantified or definite), and this labels is then used by the Anaphora Resolution Module to discard the pronoun from the list of possible referring expressions.

We intended to evaluate the system at different levels of computation so we organized an experiment related to a classical task in spoken dialogue system evaluation, distinguishing referring from non-referring pronominal expressions. This task was targeted to the sentence level processing modules: in particular, the "Pronominal Binding Module" (hence PBM) which is responsible for the search of Antecedents of pronouns in the current utterance, is fired after Quantifier Raising has been attempted. The output of the PBM is added to the DAG of current structural representation and is used by the Discourse Level processing modules to execute anaphora resolution at Discourse level[3]. In fact, what we did was to use the same setup we already organized for written text in relation to the pronoun IT, without any additional heuristics. We extended the search for "generic" non-referential pronouns to WE and YOU.

[3] For an evaluation of system's performance, see Delmonte *et al* (2006).

7.1 Experimental Results

The system has been used to parse the first 10 dialogues of the ICSI corpus for a total number of 98,523 words and 13,803 turns. This has been done to "train" the system: what happened was that, for the first 5 dialogues, we had to take care of failures. We also had to tune all the modules and procedures carefully.
In the experiment, the system has to make a two-way decision and choose to evaluate pronouns either as referential or non-referential. Here below is a table containing total values for pronouns WE/YOU/IT in all the 10 dialogues analyzed.

	Referent	Generic	Total	Found
WE	1186	706	1892	1356
YOU	1045	742	1787	1132
IT	1593	1008	2601	1627
Total	3824	2456	6280	4115

Table 3: Overall count of pronominal expressions

We had two skilled linguists to annotate pronominal WE/IT/YOU properties as either referential/non-referential. Their agreement on this task was very high with a kappa-score of 0.71. Results for the experiment are as follows

	Recall	Precision	F-Score
WE	71.67%	81.2%	76.14%
YOU	63.34%	89.3%	74.11%
IT	62.52%	84.6%	72.19%

Table 4: Results for pronominal expressions

8 Conclusions and Future Work

We have presented work carried out to extend and adapt a system for text understanding in order to make it fit for dialogue understanding. We proposed a set of expansions to cope with typical dialogue related problems, such as presence of non-sentential fragments, elliptical fragments interpretable as speech acts, massive presence of generic non-referential pronominal expressions, etc. We implemented a number of additional components: an algorithm that takes care of

overlaps and uses that information to split current utterances and temporally realign the conversational flow. A module that computes Argumentative automatic classification labels out of a small set, on top of discourse relations and other semantic markers determined by the semantic component of the system.

The system has been evaluated for two of its most important components, the newly implemented pronominal binding module and the argumentative classification module. Results are very encouraging. However, we note that in that task, labels which may cause great uncertainty and are highly ambiguous, have been lumped together to facilitate the classification task.

Of course we intend to complete the analysis of all dialogues contained in the ICSI corpus and refine our algorithms. Then, we would like to use the system with a totally different scenario, as, for instance, the Switchboard two parties dialogues and see whether the "training" carried out on the basis of multiparty dialogues may be fruitfully applied to such reduced conversational framework. In particular we still need to work at the level of DECISION labelling, and to improve the discrimination of really argumentative from pragmatically irrelevant utterance, a choice that in some cases is hard to make on an automatic basis.

APPENDIX I

EPISODES ARE MADE UP OF AGGREGATED UTTERANCES

EPISODE ISSUE No. 1
'Don'-first_1, 'Don'-first_2, 'Don'-first_3

EPISODE ISSUE No. 2
'Morgan'-first_21, 'Morgan'-'Don'-first_22

EPISODE ISSUE No. 3
'Morgan'-first_29, 'Morgan'-first_30, 'Morgan'-first_31, 'Morgan'-first_32

ARGUMENTATIVE STRUCTURE IS BUILT ON THE BASIS OF DIALOGUE ACTS AND DISCOURSE STRUCTURES
****

dial_act('Don'-nil, spact, 1, first_1, [ack, bkc, fgb, fhd])[yeah, we, had, a, long, discussion, about, how_much, ..., how, easy, we, want, to, make, it, for, people, to, bleep, things, out, '.']
EVALUATIVE CONTENT:acceptance, ARGUMENTATIVE CONTENT:[provide_expl_just(be, []), provide_expl_just(bleep_out, [agent:id18,

theme_aff:id14]), provide_expl_just(make, [agent:id18, agent:id19, theme_aff:id20, patient:id13]), provide_expl_just(want, [actor:id18, prop:id21])]
****
dial_act('Don'-nil, spact, 2, first_2, [fgb, fhd])[so, ..., '.']
EVALUATIVE CONTENT:none, ARGUMENTATIVE CONTENT:[provide_expl_just(be, [])]
****
dial_act('Don'-nil, spact, 3, first_3, [])['Morgan', wants, to, make, it, hard, '.']
EVALUATIVE CONTENT:positive, ARGUMENTATIVE CONTENT:[provide_expl_just(be, []), provide_expl_just(make, [agent:id10, theme_aff:id27]), provide_expl_just(want, [actor:id10, prop:id28])]
****
dial_act('Dave'-nil, spact, 4, first_4, [])[it, -, it, 'doesn-t_', ..., '.']
EVALUATIVE CONTENT:negative, ARGUMENTATIVE CONTENT:[disagree(do, [theme_unaff:id35]), provide_expl_just(be, [])]
****
dial_act('Don'-nil, spact, 5, first_5, [])[did, -, did, -, did, it, ..., ?]
EVALUATIVE CONTENT:suspension, ARGUMENTATIVE CONTENT:[provide_expl_just(be, []), request_expl_just(be, []), request_expl_just(do, _)]
****
dial_act('Don'-'Dave', after, 6, first_6, [])[i, 'didn-t_', even, check, yesterday, whether, it, was, moving, '.']

EVALUATIVE CONTENT:negative, ARGUMENTATIVE CONTENT:[disagree(check, [agent:id3, theme_aff:id41, prop:id43]), provide_expl_just(be, [])]
****
dial_act('Dave'-nil, spact, 7, first_7, [])[it, 'didn-t_', move, yesterday, either, when, i, started, it, '.']

EVALUATIVE CONTENT:negative, ARGUMENTATIVE CONTENT:[disagree(move, [theme_aff:id49])]
****
dial_act('Don'-'Dave', after, 8, first_8, [fgb, fhd])[so, ..., '.']
EVALUATIVE CONTENT:none, ARGUMENTATIVE CONTENT:[provide_expl_just(be, [])]
****
dial_act('Dave'-'Don', after, 9, first_9, [fgb, fhd])[so, i, 'don-t_', know, if, it, 'doesn-t_', like, both, of, us, '.']

EVALUATIVE CONTENT:negative, ARGUMENTATIVE CONTENT:[disagree(know, [actor:id2])]
****
dial_act('Don'-nil, spact, 10, first_10, [])[channel, three, ?]
EVALUATIVE CONTENT:suspension, ARGUMENTATIVE CONTENT:[request_expl_just(nil, [])]
****
dial_act('Don'-nil, spact, 11, first_11, [])[channel, three, ?]
EVALUATIVE CONTENT:suspension, ARGUMENTATIVE CONTENT:[request_expl_just(nil, [])]
****
dial_act('Dave'-'Adam', after, 12, first_12, [fhd])[you_know, i, discovered, something, yesterday, on, these, um, wireless, ones, '.']

EVALUATIVE CONTENT:negative, ARGUMENTATIVE CONTENT:[provide_expl_just(be, []), provide_expl_just(discover, [actor:id70])]
****
dial_act('Adam'-'Don', after, 13, first_13, [ack, bkc])[channel, two, '.']
EVALUATIVE CONTENT:acceptance, ARGUMENTATIVE CONTENT:[accept_expl_opin(nil, [])]
****
dial_act('Don'-nil, spact, 14, first_14, [ack, bkc])[uhhuh, ?]
EVALUATIVE CONTENT:acceptance, ARGUMENTATIVE CONTENT:[provide_expl_just(be, []), request_expl_just(be, []), request_expl_just(say, [actor:id3, theme_unaff:id86, addressee:id87])]
****
dial_act('Dave'-'Don', after, 15, first_15, [ack, bkc, fgb, fhd])[you_can_tell, if, 'it-s_', picking_up, breath, noise, and, stuff, '.']
EVALUATIVE CONTENT:acceptance, ARGUMENTATIVE CONTENT:[provide_expl_just(be, []), provide_expl_just(pick, _), provide_expl_just(pick, _), provide_expl_just(pick, _), provide_expl_just(pick, _)]
****
dial_act('Don'-nil, spact, 16, first_16, [ack, bkc, fgb, fhd])[yeah, '.']
EVALUATIVE CONTENT:acceptance, ARGUMENTATIVE CONTENT:[accept_expl_opin(nil, [])]
****
dial_act('Don'-'Dave', while, 17, first_17, [ack, bkc])[it, has, a_little, indicator, on, it, -, on, the, a_f_, '.']
EVALUATIVE CONTENT:acceptance, ARGUMENTATIVE CONTENT:[provide_expl_just(be, []), provide_expl_just(have, [experiencer:id102, tema_bound:id100])]
****
dial_act('Dave'-nil, spact, 18, first_18, [ack, bkc])[uhhuh, '.']
EVALUATIVE CONTENT:acceptance, ARGUMENTATIVE CONTENT:[accept_expl_opin(nil, [])]
****
dial_act('Dave'-nil, spact, 19, first_19, [fgb, fhd])[so, if, you, ..., '.']
EVALUATIVE CONTENT:none, ARGUMENTATIVE CONTENT:[provide_expl_just(be, [])]
****
dial_act('Dave'-nil, spact, 20, first_20, [ack, bkc, fgb, fhd])[yeah, if, you, breathe, under, -, breathe, and_then, you, see, a_f_, go_off, then, you_know, 'it-s_', ..., picking_up, your, mouth, noise, '.']
EVALUATIVE CONTENT:acceptance, ARGUMENTATIVE CONTENT:[provide_expl_just(be, []), provide_expl_just(breathe, _), provide_expl_just(go_off, _), provide_expl_just(pick_up, _)]
****
dial_act('Morgan'-nil, spact, 21, first_21, [ack, ass, bkc])[oh, 'that-s_', good, '.']
EVALUATIVE CONTENT:positive, ARGUMENTATIVE CONTENT:[accept_expl_just(be, []), provide_expl_just(be, []), provide_expl_just(be, [prop:infon434])]
****
dial_act('Morgan'-'Don', while, 22, first_22, [bkc, fgb])[because, we, have, a_lot, of, breath, noises, '.']
EVALUATIVE CONTENT:positive, ARGUMENTATIVE CONTENT:[provide_expl_just(be, []), provide_expl_just(have, [prop:infon457])]
****

dial_act('Don'-nil, spact, 23, first_23, [bkc])[yep, '.']
EVALUATIVE CONTENT:none, ARGUMENTATIVE CONTENT:[provide_expl_just(be, []), provide_expl_just(say, [actor:id3, theme_unaff:id131, addressee:id132])]
****
dial_act('Don'-nil, spact, 24, first_24, [])[test, '.']
EVALUATIVE CONTENT:none, ARGUMENTATIVE CONTENT:[provide_expl_just(be, [])]
****
dial_act('Morgan'-nil, spact, 25, first_25, [fgb])[in_fact, if, you, listen, to, just, the, channels, of, people, not, talking, 'it-s_', like, ..., '.']
EVALUATIVE CONTENT:negative, ARGUMENTATIVE CONTENT:[disagree(listen, [agent:id3, theme_aff:id77, result:id140]), provide_expl_just(be, [])]
****
dial_act('Morgan'-'Don', while, 26, first_26, [fgb])['it-s_', very, disgust, ..., '.']
EVALUATIVE CONTENT:negative, ARGUMENTATIVE CONTENT:[propose_opin_neg(be, [prop:infon535]), provide_expl_just(be, [])]
****
dial_act('Don'-nil, spact, 27, first_27, [fgb])[what, ?]
EVALUATIVE CONTENT:suspension, ARGUMENTATIVE CONTENT:[request_expl_just(nil, [])]
****
dial_act('Don'-'Morgan', while, 28, first_28, [fgb, fhd])[did, you, see, hannibal, recently, or, something, ?]
EVALUATIVE CONTENT:positive, ARGUMENTATIVE CONTENT:[provide_expl_just(be, []), request_expl_just(be, []), request_expl_just(see, [experiencer:nil, actor:id153])]
****
dial_act('Morgan'-nil, spact, 29, first_29, [fgb])[sorry, '.']
EVALUATIVE CONTENT:none, ARGUMENTATIVE CONTENT:[provide_expl_just(be, [])]
****
dial_act('Morgan'-nil, spact, 30, first_30, [ack])[exactly, '.']
EVALUATIVE CONTENT:acceptance, ARGUMENTATIVE CONTENT:[accept_expl_opin(nil, [])]
****
dial_act('Morgan'-nil, spact, 31, first_31, [fgb, neg])['it-s_', very, disconcerting, '.']
EVALUATIVE CONTENT:negative, ARGUMENTATIVE CONTENT:[provide_expl_just(be, []), provide_expl_just(be, [prop:infon610])]
****
dial_act('Morgan'-nil, spact, 32, first_32, [ack, bkc, fgb, fhd])[okay, '.']
EVALUATIVE CONTENT:acceptance, ARGUMENTATIVE CONTENT:[accept_expl_opin(nil, [])]
****
dial_act('Morgan'-'Don', after, 33, first_33, [fgb, fhd])[so, um, ..., '.']
EVALUATIVE CONTENT:none, ARGUMENTATIVE CONTENT:[provide_expl_just(be, [])]
****
dial_act('Don'-nil, spact, 34, first_34, [disf])[h, ..., '.']
EVALUATIVE CONTENT:none, ARGUMENTATIVE CONTENT:[provide_expl_just(be, [])]
****

dial_act('Morgan'-nil, spact, 35, first_35, [])[i, was, going, to, try, to, get_out_of, here, like, in, half, an, hour, '.']
EVALUATIVE CONTENT:suspension, ARGUMENTATIVE CONTENT:[provide_expl_just(be, []), suggest(get_out_of, [agent:id10, theme_aff:id180]), suggest(try, [actor:id10, prop:id181])]
****
dial_act('Morgan'-nil, spact, 36, first_36, [fgb, fhd])[um, ..., '.']
EVALUATIVE CONTENT:none, ARGUMENTATIVE CONTENT:[provide_expl_just(be, []), provide_expl_just(say, [actor:id10, theme_unaff:id187, addressee:id188])]
****
dial_act('Morgan'-nil, spact, 37, first_37, [fgb])[because, i, really, appreciate, people, coming, '.']
EVALUATIVE CONTENT:positive, ARGUMENTATIVE CONTENT:[provide_expl_just(appreciate, [agent:id10, theme_aff:id13]), provide_expl_just(be, [])

References

Allen, J., M. Dzikovska, M. Manshadi, and M. Swift (2007) Deep linguistic processing for spoken dialogue systems. In *ACL 2007*, Workshop on Deep Linguistic Processing, pp. 49-56.

Armstrong, S., A. Clark, E. Clark, L.M. Georgescul, V. Pallotta, A. Popescu-behs, M. Rajman, M. Georgescul and M. Starlander (2003) Natural language queries on natural language data: a database of meeting dialogues, in *Proceedings of NLDB'2003 conference*, Burg/Cottbus, Germany.

Bergsma, S., D. Lin and R. Goebel (2008) Distributional Identification of Non-Referential Pronouns, in *ACL-HLT 2008*, Columbus, Ohio, June 16-18, 2008, pp. 10-18.

Bresnan, J. (2000) Lexical-Functional Syntax. Oxford: Blackwell.

Bunt, H. (1979) Conversational principles in question-answer dialogues, in D. Krallmann (ed.) *Zur Theory der Frage*, Tübingen: Narr Verlag, pp. 119-141.

Clark, A. and A. Popescu-Belis (2004) Multi-level Dialogue Act Tags, In *Proceedings of SIGDIAL'04* (5th SIGdial Workshop on Discourse and Dialogue), Cambridge (MA), pp. 163-70.

ComLex:- http://nlp.cs.nyu.edu/comlex.

CoreLex:- http://www.cs.brandeis.edu/~paulb/CoreLex/ corelex.html.

Delmonte, R. (2003) Parsing Spontaneous Speech, in V. Pallotta, A. Popescu-Belis and M. Rajman (eds) *Robust Methods in Processing*

of Natural Language Dialogues, Proceedings EUROSPEECH2003, Genève: ESCA, pp. 1-6.

- (2007) *Computational Linguistic Text Processing – Logical Form, Semantic Interpretation, Discourse Relations and Question Answering*. New York: Nova Science Publishers.
- (2008) Semantic and Pragmatic Computing with GETARUNS, in J. Bos and R. Delmonte (eds.) *STEP*, London: College Pub., pp. 287-98.

Delmonte, R., A. Bristot, M.A. Piccolino Boniforti and S. Tonelli (2006) Another Evaluation of Anaphora Resolution Algorithms and a Comparison with GETARUNS' Knowledge Rich Approach, *ROMAND 2006, 11th EACL, Trento, Association for Computational Linguistics*, pp. 3-10.

Delmonte R., G. Nicolae and S. Harabagiu (2007) A Linguistically-based Approach to Detect Causality Relations in Unrestricted Text, in *Proceedings MICAI-2007*, IEEE Publications, pp. 173-85.

Galley, M., K. McKeown, E. Fosler-Lussier and H. Jing (2003) Discourse Segmentation of Multi-Party Conversation, in *Proceedings of ACL 2003*, Sapporo, Japan, pp. 562-69.

Goffman E. (1981) *Forms of Talk*. Philadelphia: University of Pennsylvania Press.

Gupta, S., M. Purver and D. Jurafsky (2007) Disambiguating Between Generic and Referential "You" in Dialog, in *Proceedings of ACL 2007 short papers*, Prague, Czech Republic.

Hillard, D., M. Ostendorf and E. Shriberg (2003) Detection of agreement vs. disagreement in meetings: Training with unlabeled data, in *Proceedings of HLT-NAACL 2003*.

Javanovich, N. and R. op den Akker (2004) Towards Adressee Identification in Multi-party dialogues, in *Proceedings of the 5th Sigdial Workshop on Discourse and Dialogue*, ACL, Pennsylvania, pp. 89-92.

Kunz W. and H.W.J. Rittel (1970) *Issues as elements of information systems. Technical Report WP-131*. Berkeley: University of California.

Lo Cascio V. (1991) *Grammatica dell'Argomentare: strategie e strutture*. Firenze: La Nuova Italia.

Mann, W.C and S.A. Thompson (1988) 'Rhetorical Structure Theory: Towards a Functional Theory Text Organization'. *Text* 8(3), pp. 243-81.

Müller, Ch. (2006) Automatic Detection on Nonreferential *It* in Spoken Multi-Party Dialog, in *Proceedings of the 11th Conference of the European Chapter of the Association for Computational Linguistics*, Trento, Italy, April 3-7, pp. 49-56.

Niekrasz J., M. Purver, J. Dowding and S. Peters (2005) Ontology-Based Discourse Understanding for a Persistent Meeting Assistant, in *Proceedings of the AAAI Spring Symposium Persistent Assistants: Living and Working with AI*. Stanford.

Pallotta, V., H. Ghorbel, A. Ballim, A. Lisowska and S. Marchand-Maillet (2004) Towards meeting information systems: Meeting knowledge management, in *Proceedings of ICEIS 2005*, Porto, Portugal, pp. 464-69.

Pallotta, V., V. Seretan and M. Ailomaa (2007) User requirements analysis for Meeting Information Retrieval based on query elicitation, in *Proceedings of the 45th Annual Meeting of the Association for Computational Linguistics (ACL 2007)*, Prague, pp. 1008-15.

Rosemberg, D. and J.A.A. Silince (1999) Common ground in computer-supported collaborative argumentation. *Paper presented at Workshop on Computer Supported Collaborative Argumentation for Learning Communities* (CSCL99).

Schegloff E. and H. Sacks (1973) 'Opening up closings'. *Semiotica* 8, pp. 289-327.

Strube, M. and C. Müller (2003) A Machine Learning Approach to Pronoun Resolution in Spoken Dialogue, in *Proceedings of the 41st Annual Meeting of the ACL*, Sapporo, Japan, 7-12 July, pp. 168-75.

Verbree, A.T. R.J. Rienks and D.K.J. Heylen, D.K.J. (2006) First Steps Towards the Automatic Construction of Argument-Diagrams from Real Discussions, in *Proceedings of the in 1st International Conference on Computat, onal Models of Argument*, 10-12 Sep, Liverpool, UK, pp. 183-94.

Normalizing speech transcriptions for Natural Language Processing

Anne Dister; Matthieu Constant; Gérald Purnelle
Université de Louvain et Facultés universitaires Saint-Louis;
Université Paris-Est; Université de Liège
anne.dister@uclouvain.be; mconstant@univ-mlv.fr; gerald.purnelle@ulg.ac.be

Abstract

Researchers in the field of spoken text processing face specific problems, all related to the nature of the data. In particular, spoken texts are full of disfluencies that constitute practical issues for automatic analysis. On the basis of a corpus of almost 500.000 words from the textual data bank of spontaneous spoken French of Valibel (http://www.uclouvain.be/valibel.html), we have especially studied four types of disfluencies: repetition, word fragments, immediate self-correction and the word euh, called "filled pause". In this paper, we show how these four types of disfluencies were automatically preprocessed in texts. The principle we used was to annotate the part of the disfluency called reparandum (according to the terminology in Shriberg, 1994), in order to keep only the repair part.

0 Introduction

Morphosyntactic tagging and syntactic parsing are key parts of Natural Language processing. Many systems now reach exploitable results for written French texts (Véronis, 2000; Clément, 2001), but there were rare attempts to automatically annotate spoken textual data (see though Mertens, 2002; Valli and Véronis, 1999). Indeed, existing software are inadequate to analyse texts transcribed from speech and face specific problems, all related to the nature of the data:

- for theoretical reasons (Blanche-Benveniste and Jeanjean, 1987), transcriptions of speech do not contain punctuation marks; nevertheless, most of the tools in Natural Language Processing are based on these marks in order to perform an initial segmentation of the text;
- texts include meta information that does not need linguistic analysis (e.g. names of speakers, information on enunciation context);
- texts contain lexical particularities specific to speech;

- finally, spoken texts are full of disfluencies, i.e. locations in the speech flow where the syntactic linearity is broken because it is interrupted for some time at a particular position on the syntagmatic axis: e.g. overlapping statements, word fragments, self-correction...

Although spoken corpus annotation does not seem to be a specific problem (Benzitoun *et al*, 2004) given the fact that there is no grammar for spoken to be opposed to a grammar for written (Blanche-Benveniste *et al*, 1990), the problems listed above need to be solved to obtain effective annotation systems. Indeed, we will see that speech transcriptions form a "new type" of texts with specificities that have to be taken into account by the analysers. In particular, disfluencies constitute a practical issue for automatic analysis of spoken texts, as many authors have already noted by reference to different languages (Adda-Decker *et al*, 2003; Bénard, 2005; Benzitoun, 2004; Benzitoun *et al*, 2004; Garside, 1995; Guénot, 2005; Nivre and Grönqvist, 2001; Oostdijk, 2003; Valli and Véronis, 1999, etc.). The specificities of spoken language considerably reduce the performance of tools initially implemented for standard written texts. The solutions adopted by the researchers in order to deal with the disfluencies are strongly dependent on the chosen approach, the task to be carried out and the tools that are being used.

Our solution consists in implementing a preprocessing module which normalizes spoken texts in order to make them compatible with standard NLP tools. On the basis of a corpus of almost 500.000 words from the textual data bank of spontaneous spoken French of VALIBEL[1] research centre, we have especially studied four types of disfluencies: repetition, word fragments, immediate self-correction and the word *euh*, called "filled pause". We have shown the regularity of these phenomena in the corpus (which are the words, part-of-speech and syntactic structures involved), and the numerous interactions between them. In this paper, we will show how these four types of disfluencies were automatically identified in texts. The principle we used was to annotate the part of the disfluency called *reparandum* (according to the terminology in Shriberg, 1994), in order to keep only the *repair* part (see below).

[1] VALIBEL for VAriétés LInguistiques du français en BELgique: http://www.uclouvain.be/valibel.html.

The paper is organized as follows. Firstly, we describe the specificities of the spoken corpus used. Then, we formalize speech particularities in order to be easily identified by the preprocessing tool. Finally, we present the resulting tool and its outputs.

1 Speech transcription

Speech transcription is not an easy task. Blanche-Benveniste and Jeanjean (1987) showed with French examples that it does not only consist for the transcribers in putting into written form what he/she hears. Transcribing requires making choices at different levels (what to be transcribed? How to transcribe?). It involves an analytic and interpretative work that has been called "translation" (Cheepen, 1995), "heuristic representation" (Mondada, 2000) or "deformation" (Bally, 1935).
From its creation in 1989, the VALIBEL research centre, which constructs and exploits large spoken corpora, established explicit transcription guidelines (Dister *et al*, 2006). They follow three main principles: use of standard spelling, no use of punctuation marks and emphasis on the speech specificities in the data (Dister and Simon, 2007).

1.1 Use of standard spelling

The transcriptions we deal with rigorously follow standard spelling conventions. Therefore, there are no graphical deformations that would consist in making a strict correspondence with the pronunciation[2], as it can be frequently found in noble dialogues. From a spelling perspective, transcriptions cannot be distinguished from standard written French: no wild elisions (e.g. *j'suis* standing for *je suis* (*I am*), *p'tit* for *petit* (*little*)), no graphical "monsters" (e.g. *ché pas* pour *je sais pas* (*I don't know*), *pasque* for *parce que* (*because*)). All lexemes used can be found in standard texts as listed in language references like dictionaries. From a Natural Language Processing perspective, words are analysed on the basis of lexical resources containing them.

[2] When necessary, transcribers can add pronunciation information in meta-tags.

1.2 No punctuation marks

Usually, corpora of speech transcriptions built for linguistic research do not contain punctuation marks. Indeed, there exist no strict correspondences between prosodic phenomena and written punctuation. A short pause does not always correspond to a comma in written texts. Furthermore, a longer pause does not systematically imply the use of stronger punctuation marks.

Blanche-Benveniste and Jeanjean (1987: 139) plead for speech transcriptions with no punctuation marks. They argue that their use implies that transcribers suggest an analysis before having performed it. This is why the notion of sentence has been abandoned in the studies on speech production. Therefore, we might wonder what minimal unit is required by NLP tools that need sentence segmentation.

Although punctuation marks are not used to annotate the corpus, silent pauses are inserted to help reading. There are three degrees of pause that were subjectively assigned depending their duration: / (short pause), // (long pause), /// (silence).

1.3 Emphasis on spoken specificities

Disfluencies

Studies on spoken language highlighted specific phenomena that are generally called *disfluencies*. They correspond to locations of the speech flow where the linearity is broken because it stops for some time on the syntagmatic axis. We name this way punctuation words (*ben*, *bon*...), the filled pause *euh* (*uh*), repetition of words or word sequences (cf. 2.1), immediate self-corrections such as *le la fille* (*the the girl*, cf. 2.2), word fragments (transcribed with the slash symbol /: *à Bru/à Bruxelles* (*in Bru/in Brussels*), cf. 2.3), etc.

Disfluency transcription demands a careful attention from transcribers, in order to write down such phenomena that are usually filtered by an ordinary listening. Indeed, they are so frequent in spontaneous speech that we unconsciously tend to ignore these marks.

Speaking slots

The sound continuum, that has become linear with the transcription, is divided into speaking slots, defined by the change of speaker. In our transcriptions, the sequencing of speaking slots is presented

horizontally: words of the speakers succeed to each other top-down on the screen. Each paragraph represents the intervention of a speaker.
This organisation that Edwards (1995) calls vertical format is coherent with our reading habits: we start reading from the top of the screen and what we read before occurs in time before we read after. Theatre texts adopted this format from a long time.

Overlapping statements

In standard spontaneous conversations, it is very frequent that two or more persons speak at the same time. We therefore have overlapping speech statements. In our transcription convention, symbols | and - delimit overlapping segments (|- for the beginning of the overlapping and -| for the end). For instance, in the following transcription

L1 je le connais |- depuis longtemps
L2 oui tu -| l'avais rencontré à mon mariage
L1 I know him |- for a long time
L2 yes you -| had met him at my wedding

speaker L2 starts to speak while a speaker L1 is already speaking; L2 continues and L1 stops.
However, it can happen than the second speaker starts to speak during the first speaker' speech, but the latter keeps on speaking after the overlap. We then have an internal overlapping segment that is transcribed as following:

L1 je l'aime |- vraiment beaucoup <L2> je sais -| ce chercheur

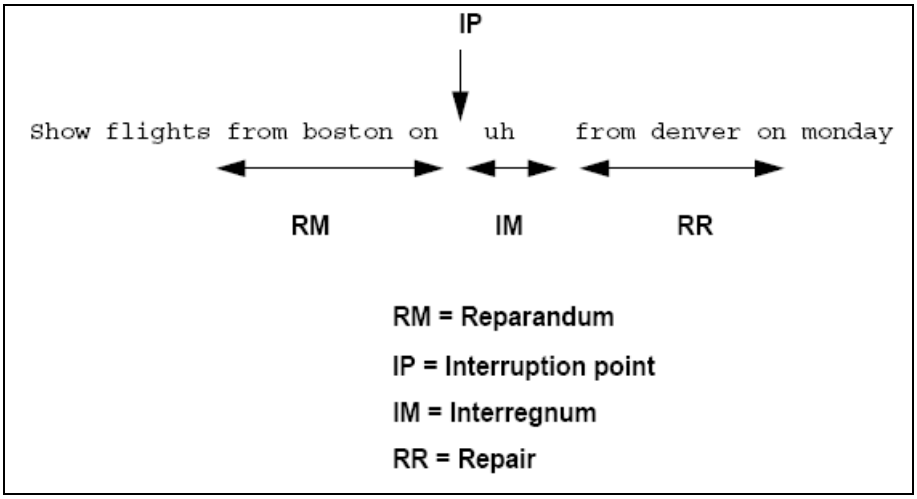
1.4 Transcription example

ileGF0 une une trémie / ça veut dire quoi
ilePA2 une trémie justement une trémie i/ |- c'est une < ileGF0> oui -| un tunnel une trémie chez nous c'est / c'est le c'est c'est ce qu'on appelle un tunnel
ileGF0 ah d'accord
ilePA2 hein |- mais < ileGF0> mm -| une pet/ un petit tunnel qui n'est pas très long
ileGF0 mm
ilePA2 or une trémie euh grammaticalement c'est une chose qui s'en/ qui s'enfonce plutôt dans la terre
ileGF0 a a hopper / what does it mean
ilePA2 a hopper precisely a hopper i/ |- it's a < ileGF0> yes -| a tunnel a hopper in our country it's / it's the it's it's what is called a tunnel
ileGF0 oh well !
ilePA2 eh |- but < ileGF0> mm -| a lit/ a little tunnel which is not very long

ileGF0 mm
ilePA2 now a hopper er grammatically it's a thing which pen/ penetrates in the earth

2 Identifying disfluencies

Shriberg (1994: 7-9), following Levelt (1989), represented the disfluent sequence by splitting it into four distinct elements corresponding to three regions:



• ***reparandum***: the *reparandum* (*RM*) is the part produced by the speaker that will not be kept and that will be replaced later by the *repair*;
• ***interrupting point***: the interrupting point (IP) is the moment of the speech that coincides with the end of the *reparandum*. It has an empty textual content;
• ***interregnum***[3]: the *interregnum* (*IM*) is the region that begins at the end of the reparandum and ends at the begining of the repair. It can contain an *editing term*, i.e. a silent pause, a filled pause, or several attempts of unachieved reformulation;
• ***repair***: the *repair* (*RR*) indicates the correction of the *reparandum*.

[3] The *interregnum* corresponds to Levelt's editing phrase.

For a long time, researchers showed regularities constraining disfluencies, especially in English (Blankenship and Kay, 1964; Cook, 1971). This regularity also observed in our data will allow us to formalize effectively this phenomena[4] and automatically identify them during the preprocess stage.
For us, the preprocessing stage for disfluencies consists in identifying the reparandum and the repair. At the final state of the process, the reparandum would be assigned a specific tag in order to only take into account the repair part in further automatic analyses (e.g. morphosyntactic tagging, chunking, etc.).

2.1 Repetitions

A repetition is a sequence of two (or more) contiguous graphically identical forms. The identical forms can be words or groups of words such in the example below involving repetitions of words *sans* (without) and *la* (the):

ilrMS1 je sais pas / parler sans accent pour moi c'est sans // sans // sans bafouiller sans / sans sans se tromper de mots quoi sans sans sans que la la langue fourche quoi [ilrMS1r]
ilrMS1 I don't know / to speak without an accent for me it's without // without // stamming without / without without getting the wrong word what without without without a slip of the tongue what

Repetitions temporary break the linearity of the statement, by staying on the same location of the syntagmatic axis. The grid representation proposed by Blanche-Benveniste *et al* (1979) allows for taking account of the phenomenom. It superposes repeated terms:

ilrMS1 je sais pas / parler sans accent pour moi
c'est sans
// sans
// sans bafouiller sans
/ sans
sans se tromper de mots quoi sans
sans
sans que la

[4] It is based on a systematic linguistic study of the disfluencies and silent pause marks occurring in a 440.000-word corpus (around 40 hours of speech). For more details, see Dister, 2007.

la langue fourche
quoi

ilrMS1 je sais pas / parler sans accent pour moi c'est **{sans // sans //, .IGN+rep }** **sans** bafouiller **{sans / sans, .IGN+rep }** **sans** se tromper de mots quoi **{sans sans, .IGN+rep }** **sans** que **{la,.IGN+rep}** **la** langue fourche quoi

The reparandum plus interregnum parts are tagged between curly brackets with the tag IGN+rep (IGN for ignore and rep for repetition).

2.2 Immediate self-correction

Immediate self-correction phenomena are variants of repetition ones. In self-correction, one of the morphosyntactic features of the repeated element varies, as it is shown below

ileFN1 et le journalisme et puis euh **le** les études de journalisme en soi ne me plaisaient pas [ileFN1r]
and I did not like journalism and er studying journalism in itself
In the example, *les* is the plural form of *le*.
The tagging is the same as the one for repetitions, except the tag *cor* for *correction*.
ileFN1 et le journalisme et puis euh **{le,.IGN+cor}** **les** études de journalisme en soi ne me plaisaient pas [ileFN1r]

2.3 Word fragments

A word fragment consists of an interruption of the morpheme being enunciated. According to the terminology in Pallaud (2002), word fragments can be divided into three categories: completed word fragments, corrected word fragments and unachieved word fragments. The three cases are illustrated respectively in the following statements:

iljDV1 apprendre ça c'est transm/ transmettre un savoir donc ça c'est apprendre communiquer euh euh (…) [iljDV1r]
to teach this is to pass a knowledge on so this is to teach to communicate er er
accFJ1 (…) j'ai été à plusieurs reprises avec mes parents en Auvergne je trouvais aussi qu'ils avaient aussi un accent qui était pas mal euh // typique / par contre les J/ les Bretons j'ai jamais su / jamais vu qu'ils avaient d'accent moi (…) [accFJ1r]
I've been on many occasions with my parents in the Auvergne so I found they had too an accent which was quite er // typical / on the other hand / Bretons I've never known / never seen they had an accent I (...)

ilrVI2 m quand un néerlandophone parle français / euh je trouve que ça ne fait pas bien du tout / par rapport à quelqu'un qui parle bien fran/ comme un Bruxellois par exemple (…) [ilrVI2r]
m when a Dutch speaker speaks French / er I find it doesn't make it at all / in comparison with somebody who speaks good Fren/ like a person from Brussels for instance (...)

The two first types of word fragments are subject to the same type of annotation as repetitions and immediate self-correction (the tag is *frag*). The unachieved word fragments are also annotated but without the repair part.

3 Text segmentation

The prepocessing is not only limited to the tagging of the disfluencies. It also requires a new segmentation of the texts, that consists in extracting internal overlapping segments and segmenting speaking slots into smaller parts.

3.1 Speaking slots and overlapping segments

Like disfluencies, overlapping markers break the linearity of the reading. We though observed that in almost all cases, a speech overlapping is not a syntactic break of the statement: the speaker being overlapped continues speaking as if he/she were not interrupted. In the preprocessing stage, we annotate the starting and ending markers of the overlapping segments with the tags IGN+over and.IGN+overEnd. The speaking slots are identified by unique numbers (e.g. #245). For internal overlapping fragments, we extract them in the form of a new speaking slot being referred by the overlapped speech fragment (e.g. @246). Both examples given in section 1.3 are respectively transformed as follows:

{#123,.IGN+slot} {L1,.IGN+speaker} je le connais {|-,.IGN+over} depuis longtemps
{#124,.IGN+slot} {L2,.IGN+speaker} oui tu {-|,.IGN+overEnd} l'avais rencontré à mon mariage
{#245,.IGN+slot} {L1,.IGN+speaker} je l'aime {|-,.IGN+over} vraiment vraiment beaucoup @246{-|,.IGN+overEnd} ce chercheur
{#246,.IGN+slot} {L1,.IGN+speaker}je sais

3.2 Sentence segmentation

Traditionally, Natural Language Processing tools work on the sentence level. Therefore, the first task to do is to segment the graphical chain into tokens (roughly speaking words) and in sentences. Nevertheless, our transcriptions do not contain any punctuation marks and the only a priori segmentation available is the one in speaking slots. As some slots are very long, it is necessary to cut the text into smaller units. For this, we examined whether silent pauses could be the basis of a relevant initial segmentation for automatic annotation. Following studies by Duez (1991) and Candea (2000), we made the hypothesis that transcribers put silent pauses at preferential location in terms of syntactic structure of the text, allowing for relevant regrouping for automatic analysis. In practice, it seems that long pause and silence marks are good candidates for text segmentation in smaller units corresponding roughly to chunks (Abney, 1991), with a low error rate (see Dister, 2008 for further details of the analysis).

4 Outputs

The preprocessing part handles other phenomena in spoken texts such as the speaker identification, the tagging of *euh* (uh), phonetic or paralinguistic markers, etc. The principle is the same: tagging text portions with tag IGN indicating to the analyzer that it has to ignore it. The sample given in section 1.4 is then transformed by our tool in:

{S}{#1,.IGN+slot}{ileGF0,.IGN+speaker} {une,.IGN+rep} une trémie {/,.IGN+meta} ça veut dire quoi {S}
{S}{#2,.IGN+slot} {ilePA2,.IGN+speaker} une trémie justement une trémie {i/,.IGN+frag} {|-@3,.IGN+over} c'est {une,.IGN+corr} {-|,.IGN+overEnd} un tunnel une trémie chez nous {c'est,.IGN+rep} {/,.IGN+meta} c'est le {c'est,.IGN+rep} c'est ce qu'on appelle un tunnel {S}
{S}{#3,.IGN+slot} { ileGF0,.IGN+speaker} oui {S}
{S}{#4,.IGN+slot} {ileGF0,.IGN+speaker} ah d'accord {S}
{S}{#5,.IGN+slot} {ilePA2,.IGN+speaker} hein {|-@6,.IGN+over} mais {-|,.IGN+overEnd} {une pet/,.IGN+frag} un petit tunnel qui n'est pas très long {S}
{S}{#6,.IGN+slot} {ileGF0,.IGN+speaker} mm {S}
{S}{#7,.IGN+slot} {ileGF0,.IGN+speaker} mm {S}

{S}{#8,.IGN+slot} {ilePA2,.IGN+speaker} or une trémie {euh,.IGN+euh} grammaticalement c'est une chose {qui s'en/,.IGN+frag} qui s'enfonce plutôt dans la terre {S}

5 Conclusions

The specificities of texts transcribed from speech rise problems for syntactic and morphosyntactic analyzers. In this paper, we described a preprocessing tool allowing for handling the difficulties inherent to speech transcriptions. The good results obtained by a morphosyntactic tagger (Dister, 2007) and a chunker chunks (Blanc *et al*, 2008) show the relevancy of our choices.

References

Abney, S. (1991) Parsing by chunks, in R. Berwick, S. Abney and C. Tenny (eds) *Principle-based parsing: Computation and Psycholinguistics*, Boston: Kluwer Academic Publishers, pp. 257-78.

Adda-Decker, M., B. Habert, C. Barras, G. Adda, P. Boula de Mareüil and P. Patrick (2003) 'A disfluency study for cleaning spontaneous speech automatic transcripts and improving speech language models', *ISCA Tutorial and Research Workshop on Disfluency in Spontaneous Speech (DiSS'03)*, Gothenburg: University of Gothenburg, pp. 67-70.

Bally, C. (1935) *Le Langage et la Vie*. Zurich: Max Niehans (2nd ed.).

Bénard, F. (2005) *Normalisation de corpus oraux: des métadonnées à l'annotation des transcriptions*. Université Paris-3, Sorbonne Nouvelle, Mémoire de maitrise.

Benzitoun, C. (2004) 'L'annotation syntaxique de corpus oraux constitue-t-elle un problème spécifique ?', *Actes de RÉCITAL* (21 avril 2004, Fès), pp. 12-22.

Benzitoun, C., E. Campione, J. Deulofeu, S. Henry, F. Sabio, S. Teston, A. Valli and J. Véronis (2004) 'L'analyse syntaxique de l'oral : problèmes et méthode'. *Journée d'étude de l'ATALA sur l'annotation syntaxique de corpus* (15 mai 2004, Paris), pp. 13-22.

Blanc, O., A. Dister, M. Constant, and P. Watrin (2008) 'Corpus oraux et chunking', *Actes des 27es Journées d'étude sur la parole (JEP 2008)*, Avignon, 9-13 juin 2008.

Blanche-Benveniste, C., B. Borel, J. Deulofeu, J. Durand, A. Giacomi, C. Loufrani, B. Meziane and N. Pazery (1979) 'Des grilles pour le français parlé'. *Recherches sur le français parlé* 2, Université de Provence, pp. 163-205.

Blanche-Benveniste, C. and C. Jeanjean (1987) *Le Français parlé. Transcription et édition*. Paris: Didier Érudition.

Blanche-Benveniste, C., M. Bilger, C. Rouget and K. van den Eynde (1990) *Le Français parlé. Études grammaticales*. Paris: CNRS Éditions.

Blankenship, J. and C. Kay (1964) 'Hesitation phenomena in English Speech: a study in distribution'. *Word* 20, pp. 360-72.

Candea, M. (2000) *Contribution à l'étude des pauses silencieuses et des phénomènes dits «d'hésitation» en français oral spontané*. Université de Paris 3 – Sorbonne nouvelle, not published dissertation.

Cheepen, C. (1995) Discourse considerations in transcription and analysis, in G. Leech, G. Myers and J. Thomas (eds) *Spoken English on Computer. Transcription, Mark-up and Application*, New York: Longman, pp. 135-43.

Clément, L. (2001) *Construction et exploitation d'un corpus syntaxiquement annoté pour le français*. Université Paris 7, not published dissertation.

Cook, M. (1971) 'The Incidence of Filled Pauses in Relation to Part of Speech'. *Language and Speech* 14, pp. 135-50.

Dister, A. (2007) *De la transcription à l'étiquetage morphosyntaxique. Le cas de la banque de données textuelles orales Valibel.* Université de Louvain, not published dissertation.

– (2008) La notation subjective de la pause constitue-t-elle un bon indice pour le découpage de corpus oraux?, in M. Constant, A. Dister, L. Emirkanian and S. Piron (eds) *Description linguistique pour le traitement automatique du français*, Cahiers du Cental 5, Louvain-la-Neuve: Presses universitaires de Louvain, pp. 165-86.

Dister, A., M. Francard, G. Geron, V. Giroul, P. Hambye, A.C. Simon and R. Wilmet (2006) *Conventions de transcription régissant les corpus de la banque de données Valibel.* http://valibel.fltr.ucl.ac.be, corpus oraux, conventions de transcription.

Dister, A. and A.C. Simon (2007) 'La transcription synchronisée des corpus oraux. Un aller-retour entre théorie, méthodologie et traitement informatisé'. *Corpus and text linguistics in Romance*

languages, Arena Romanistica 1/1, Bergen: Presses de l'Université de Bergen, pp. 54-78.
Duez, D. (1991) *La Pause dans la parole de l'homme politique*. Paris: Éditions du CNRS.
Edwards, J.A. (1995) Principles and alternative systems in the transcription, coding an mark-up of spoken discourse, in G. Leech, G. Myers and T. Jenny (eds) *Spoken English on computer. Transcription, mark-up and application*, New York: Longman, pp. 19-34.
Garside, R. (1995) Grammatical tagging of the spoken part of the British National Corpus: a progress report, in G. Leech, G. Myers and J. Thomas (eds) *Spoken English on Computer. Transcription, Mark-up and Application*, New York: Longman, pp. 161-7.
Guénot, M.-L. (2005) 'Parsing de l'oral: traiter les disfluences'. *Actes de TALN 2005* (6-10 juin, Dourdan), pp. 323-332.
Levelt, W.J.M. (1989) *Speaking: from intention to articulation*. Cambridge: MIT Press.
Mertens, P. (2002) Les corpus de français parlés ELICOP : consultation et exploitation, in J. Binon, P. Desmet, J. Elen, P. Mertens and L. Sercu (eds) *Tableaux Vivants. Opstellen over taal-en-onderwijs aangeboden aan Mark Debrock*, Leuven: Universitaire Pers, pp. 101-116.
Mondada, L. (2000) 'Les effets théoriques des pratiques de transcription'. *LINX* 42, revue de l'Université de Paris X-Nanterre, pp. 131-50.
Nivre, J. and L. Grönqvist (2001) 'Tagging a Corpus of Spoken Swedish'. *International Journal of Corpus Linguistics* 6, 1, pp. 47-78.
Oostdijk, N. (2003) Normalization and disfluencies in spoken language data, in S. Granger and St. Petch-Tyson (eds) *Extending the scope of corpus-based research. New applications, new challenges*, Amsterdam-New York: Rodopi, pp. 59-70.
Pallaud, Be. (2002) 'Les amorces de mots comme faits autonymiques en langage oral'. *Recherches sur le français parlé* 17, Université de Provence, pp. 79-101.
Shriberg, E. (1994) *Preliminaries to a Theory of Speech Disfluencies*. Université de Berkeley, not published dissertation.
Valli, A. and J. Véronis (1999) 'Étiquetage grammatical des corpus de parole: problèmes et perspectives'. *Revue française de linguistique*

appliquée 4, 2, pp. 113-33.
Véronis, J. (2000) Annotation automatique de corpus: panorama et état de la technique, in J.-M. Pierrel (ed.) *Ingénierie des langues*, Paris: Hermès, pp. 111-29.

È possibile "guardare la comunicazione?"

Esperienze di dialogo tra Analisi della Conversazione e video osservazione dei processi comunicativi

Federico Farini
Università di Modena e Reggio Emilia, Dipartimento di Scienze del Linguaggio e della Cultura
federico.farini@unimore.it

Abstract

This article discusses some issues related to the combined use of Conversation Analysis as a method for sociological research on interaction, and videotaping as a technique for data production. These considerations rely on research experiences at the Department of Language and Culture Sciences (University of Modena and Reggio Emilia, Italy) where, since 2005, a team composed of both sociologists and linguists cooperate in various research projects on interaction at the intersection between linguistics, sociolinguistics and sociology.
Conversation Analysis is a methodology devoted to the analysis of interaction, meaning with "interaction" a sequence of actions which have social import, bringing about reactions from co-participants.
Among the basic theoretical presuppositions of Conversation Analysis we have the concept of "sequentiality": any action, in any social situation, is retrospectively connected to the sequence of action preceding it, and prospectively connected to following actions. According to this, the article discusses the issue of videotaping as a relevant technique to produce real-time recordings which impart to the researcher the ability to analyze the connections between actions in sequences as they naturally occur. Thus, videotaping is the only technique which makes it available for scientific observation the delicate complexity of interaction, making it possible for Conversation Analysis to recognize structures of social action in a deductive way.

1 L'Analisi della Conversazione, una metodologia di ricerca sociale

L'Analisi della Conversazione (AC) è una metodologia di ricerca qualitativa che rende possibile analizzare empiricamente i presupposti sociali e culturali dell'interazione umana nei più diversi contesti sociali.
Il fondamento epistemologico dell'AC è il seguente: la presenza di connessioni sistematiche tra azioni, per cui al verificarsi di una certa azione ci si aspetta il verificarsi di un determinato tipo di azione

successiva, rappresenta una traccia della presenza di strutture efficaci nel coordinare in forme stabili, quindi, reciprocamente prevedibili, il comportamento dei partecipanti all'interazione. Pertanto, l'AC è una metodologia adatta a ricerche sociologiche che non si accontentano di prescrivere induttivamente l'esistenza di strutture fondamentali dell'azione sociale, ma vogliono verificarla empiricamente.
Nel momento in cui la ricerca sociologica si pone il problema del rapporto tra azione e struttura sociale, essa non può che scegliere come proprio "campo" il luogo dove questo rapporto diventa empiricamente osservabile: l'interazione.
L'interazione è una situazione sociale caratterizzata dalla compresenza dei partecipanti, che può essere corporea o solo vocale. Chiunque prenda parte a un'interazione non può non essere influenzato dalla consapevolezza di essere *presente-con* (Luhmann, 1984; Luhmann and De Giorgi, 1992). Nell'interazione ogni partecipante è conscio di almeno due cose: la prima è che mentre osserva i propri interlocutori, a sua volta è oggetto di osservazione; la seconda è che anche i suoi interlocutori sanno di essere osservati. Chiunque prenda parte a un'interazione tiene conto di questa duplice condizione di osservatore/osservato.
Nonostante sia il medium fondamentale dell'interazione, il linguaggio *in sé* non rappresenta l'oggetto di studio dell'AC. A differenza della sociolinguistica, l'AC non vuole portare alla luce l'influenza di macrostrutture sociali intese come variabili indipendenti (ad es., potere, genere, classe) sulle strutture del linguaggio. L'AC guarda al linguaggio come medium per l'azione, e la comprensione dei significati dell'azione (Hutchby and Woffitt, 2008). In questa prospettiva, il linguaggio non solo rispecchia e riproduce strutture sociali, ma può creare i presupposti della loro evoluzione (Baraldi and Gavioli, 2007).

2 I principi fondamentali dell'Analisi della Conversazione

Nella sezione precedente abbiamo definito l'AC come importante risorsa metodologica per un'osservazione empirica delle strutture dell'azione sociale. Per il ricercatore sociale è però importante interrogarsi su quello che concretamente fa, quando utilizza l'AC come metodo per fare ricerca, in questo modo applicando un principio epistemologico fondamentale per l'AC, quello di "problematizzare la

routine" (Garfinkel, 1967), anche alla sua pratica di ricerca. Riflettere su quello che si fa, quando si fa ricerca, rappresenta un'operazione molto importante, poiché solo se la ricerca diventa a sua volta oggetto di ricerca (Spencer-Brown, 1969) le sue procedure possono essere sviluppate, corrette, migliorate.
Iniziamo con il riflettere su quello che comporta adottare l'AC come metodo di ricerca sociologica. Utilizzare l'AC per analizzare l'interazione obbliga il ricercatore ad appropriarsi di tre principi fondamentali:

1) *Naturalismo*. L'osservazione dell'interazione deve avere per oggetto interazioni reali e naturali;
2) *Sequenzialità*. Il significato dell'azione sociale è dato dalla sua posizione in una sequenza di azioni, dal suo essere collocata prima e dopo qualcosa d'altro;
3) *Adiacenza*. I legami tra azioni adiacenti rappresentano la cornice normativa fondamentale dell'ordine sequenziale dell'interazione.

2.1 L'osservazione naturalistica dell'interazione

Oggetto dell'AC sono interazioni *reali* e *naturali*, cioè non inventate dal ricercatore e non costruite artificialmente con fini sperimentali. Il principio dell'osservazione naturalistica dell'interazione, per cui l'AC discute scambi *reali* e non esempi inventati dal ricercatore a scopo argomentativo, differenzia il "fare AC" dal "fare filosofia del linguaggio".
Il *naturalismo* differenzia il "fare AC" dal "fare etnometodologia". L'etnometodologia inventò il principio di problematizzazione della *routine* sociale eleggendo la quotidianità come oggetto di studio, ed è per questo considerata il fondamento dell'AC (Atkinson and Heritage, 1984; Giglioli and Fele, 2000); rispetto all'AC, però, utilizza una metodologa di ricerca differente, basata sulla procedura del *breaching experiment* (Garfinkel, 1967) che consiste nella produzione sperimentale di interazioni, dove ricercatori addestrati violano le convenzioni sociali deducendo la loro rilevanza come strutture dell'azione sociale in base al tipo ed all'intensità delle reazioni degli interlocutori. A differenza dell'AC, quindi, l'etnometodologia lavora su interazioni reali, ma non naturali.

2.2 Sequenzialità

Per l'AC ogni azione sociale acquista significato all'interno di sequenze di azioni. Dal punto di vista sociologico, il concetto di sequenzialità ha una valenza che travalica quello di serialità. Per l'AC ogni azione sociale non solo e non tanto è seriale perché avviene dopo e prima di qualcos'altro ma è sequenziale, poiché acquisisce significati sociali diversi, all'interno di diverse sequenze di azioni interconnesse.
Nell'interazione ogni azione è legata retrospettivamente a quello che è avvenuto prima, che ne rappresenta il contesto locale, e prospettivamente a quello che avviene in seguito, che risulta esserne inevitabilmente condizionato. A partire dall'osservazione naturalistica dell'interazione, l'applicazione del principio di sequenzialità accende i riflettori sulle connessioni tra azioni. Per il sociologo empirico, la presenza di regolarità in queste connessioni è un argomento, genuinamente deduttivo, a favore della presenza di efficaci strutture dell'azione sociale.

2.3 Adiacenza

Conseguentemente al principio di sequenzialità, l'AC non si interessa a singoli atti linguistici, ma alle sequenze di azioni che si sviluppano, retrospettivamente e prospettivamente, nel tempo. La forma basilare di sequenza è la *coppia adiacente* di enunciati (Schegloff and Sacks, 1973), dove la prima parte richiede una particolare seconda, o una ristretta gamma di seconde parti.
Esempi di coppie adiacenti sono le sequenze "saluto-saluto", "domanda-risposta", "valutazione-accordo". Il rapporto tra le parti delle coppie adiacenti è descritto dal concetto di *rilevanza condizionale* (ivi): se in seguito ad una prima parte di coppia adiacente non è prodotta la seconda parte attesa, allora chi ha prodotto la prima parte deve spiegarsi quest'assenza, e reagirvi in qualche modo.
Osservare il grado di efficacia delle prime parti di coppie adiacenti nell'orientare l'azione dei partecipanti, e la reazione dei loro autori ad eventuali insuccessi, permette di dedurre l'orientamento dei partecipanti a ruoli ed aspettative generalizzate, ossia alle strutture dell'azione sociale. Le seconde parti delle coppie adiacenti sono un

luogo fondamentale per l'interazione perché mostrano se e come le azioni precedenti sono state comprese. La rilevanza delle seconde parti delle coppie adiacenti travalica gli aspetti linguistico-cognitivi, coinvolgendo la dimensione sociale. La scelta di fornire, ovvero di non fornire, la "seconda parte attesa" dice molto sulla volontà di affiliazione o dis-affiliazione verso l'autore della prima parte della coppia, e lo dice sia a chi partecipa all'interazione sia al sociologo che analizza l'interazione.

3 Dall'interazione "reale e naturale" ai dati scientifici

Se si considera l'AC una risorsa metodologica fondamentale per l'osservazione empirica delle strutture dell'azione sociale, allora è necessario rispondere a due interrogativi: quali sono i dati su cui lavora l'AC? Com'è possibile produrli?
L'AC lavora su trascrizioni *integrali* e *accurate* di tutto quello che accade in un'interazione. Queste trascrizioni devono riportare tutto quello che accade comprese pause, sovrapposizioni tra parlanti e altri elementi sonori (applausi, risate, sospiri). Qualora si disponga di registrazioni video, a essere riportate sono anche le azioni non verbali (annuire, scuotere il capo, indicare qualcuno, gli sguardi). L'accuratezza nel riportare tutto quello che è accaduto nell'interazione, il modo in cui è accaduto, e il momento in cui è accaduto è un vero e proprio imperativo epistemologico per l'AC.
Innalzando a oggetto di ricerca pratiche interazionali apparentemente banali, che però rappresentano l'enorme mole di lavoro "dietro le quinte" che consente il prodursi di ogni forma di socialità, l'AC non può scartare niente di quello che accade all'interno di un'interazione come irrilevante, ma anzi fa degli scambi internazionali quotidiani il proprio oggetto di studio.
La sociologia si è impegnata per decenni nel "rendere normale lo strano", nello spiegare quello che appare "deviare" rispetto a una "normalità". Di per sé, la normalità ha sempre riscosso ben poco interesse, tradizionalmente considerata uno "sfondo" da cui si stagliano azioni devianti e grandi processi sociali di trasformazione.
In modo del tutto opposto l'AC (e la ricerca sociologica che la utilizza come metodo) si impegna a "rendere strano il normale", a fare della normalità oggetto di ricerca sociologica. I processi sociali che la sociologia classica ha sempre ritenuto "ovvi", le attività quotidiane

che si riproducono senza apparenti problemi, la serie infinita di episodi di "normalità" sono per l'AC tutto fuorché banali. "Fare le cose normali" è considerato dall'AC come esito di un incessante e delicato lavoro sociale, dove ogni soggetto contribuisce con le sue azioni a riprodurre forme ordinate di socialità. Pertanto, la "normalità" è considerata dall'AC non solo meritevole di attenzione sociologica, ma oggetto di studio per eccellenza.

Nel momento in cui adotta l'AC come metodo, la ricerca empirica sulle strutture dell'azione sociale si abilita ad appropriarsi di una miniera di dati potenzialmente illimitata, che coincide con l'insieme delle ordinarie attività della quotidianità. Il problema è trasformare queste attività, che ovviamente accadono a prescindere dall'interesse che riscuotono presso i sociologi, in dati scientifici.

Per fare questo, c'è bisogno di trascrizioni che riportano sulla pagina non solo quello che accade nell'interazione, ma "quello che accade, chi lo fa, quando lo fa e come lo fa", nel più fine dettaglio possibile. Non si tratta solo di riportare le azioni verbali e non verbali, ma anche il punto preciso in cui compaiono, preservando i legami retrospettivi e prospettici con quello che precede e con quello che segue.

Per produrre trascrizioni tanto raffinate c'è bisogno di una tecnica per la registrazione dell'interazione che ne preservi con il massimo grado di dettaglio le componenti verbali, quelle non verbali, l'ordine in cui esse si presentano e la natura dei loro legami sequenziali. Fin dai suoi esordi, l'AC riconosce come dati accettabili esclusivamente registrazioni su supporto magnetico (poi anche digitale) di interazioni naturali (Atkinson and Heritage, 1984; Sacks, 1992). Del resto, i principi metodologici del naturalismo, della sequenzialità e dell'adiacenza (vedi 2) sono applicabili solo a interazioni registrate in tempo reale, che preservano la temporalità delle interazioni e ne permettono un'analisi ripetuta.

L'osservazione ripetuta di interazioni reali e naturali può portare alla luce il grande lavoro sociale richiesto da pratiche quotidiane che, nonostante la scarsa attenzione che la ricerca sociologica rivolge loro, rappresentano comunque risorse fondamentali della socialità umana. Una domanda interessante, quindi, potrebbe essere: quali tecniche sono disponibili per registrare e fissare su un supporto resistente, magnetico o digitale, le interazioni?

4 Le tecniche di produzione dei dati che servono a una ricerca sociologica sull'interazione

Per rispondere a questo interrogativo, è per noi possibile richiamare i risultati di una riflessione metodologica maturata nel corso di una pluriennale attività di ricerca sull'interazione, condotta da un'équipe integrata di sociologi e linguisti afferenti al Dipartimento di Scienze del Linguaggio e della Cultura, presso l'Università di Modena e Reggio Emilia (Baraldi, 2005; Iervese, 2006; Baraldi, 2008; Farini, 2009).
Si intende riflettere sull'utilizzo combinato dell'AC come metodologia di analisi dell'interazione, e della video-osservazione (VO) come tecnica di produzione dei dati, così come questa équipe sta sperimentando da anni. Nelle sezioni precedenti abbiamo discusso della rilevanza dell'AC per la ricerca sociale (1), abbiamo presentato i principi fondamentali dell'AC (2), abbiamo descritto le caratteristiche dei dati su cui lavora l'AC (3); è ora il momento di considerare la tecnica che riteniamo più adatta a preservare la complessità e la delicatezza dell'interazione: la registrazione audio e video dell'interazione. La VO non è certo l'unica tecnica di produzione dei dati disponibile per chi si interessa allo studio dell'organizzazione del sociale. Riteniamo però che sia la tecnica più efficace, come argomenteremo anche attraverso la comparazione con altre tecniche disponibili.

4.1 L'osservazione partecipante e le note di campo

Partiamo dalla tecnica generalmente associata alla metodologia etnografica dell'osservazione partecipante, che potremmo definire "tecnica delle note di campo" e che prevede la presenza di un ricercatore sul campo, chiamato a registrare i fenomeni sociali più rilevanti su supporto cartaceo, ad esempio un taccuino.
La tecnica delle note di campo offre un indiscutibile vantaggio nella limitata intrusività del mezzo di registrazione-dati (corpo del ricercatore+penna+supporto cartaceo). Essa non è però in grado di produrre dati utili per un'analisi dell'interazione interessata a osservare come i presupposti sociali e culturali della socialità si manifestano in forme d'ordine sequenziale, le quali sono troppo delicate perché siano percepite in presa diretta, e troppo complesse per

essere riprodotte fedelmente in forma di note sul taccuino (Fele, 2007).
Si prenda, ad esempio, l'interazione educativa; una sfida per il ricercatore può consistere nello scoprire se e come la relazione sociale "educatore-educando" si manifesti e si riproduca empiricamente in sequenze di azioni e re-azioni. Diverse ricerche che utilizzano l'AC come metodologia di ricerca evidenziano che nell'interazione educativa una sequenza di azioni molto comune è una tripletta di turni del tipo: 1) turno di parola dell'educatore, 2) annuire degli educandi, 3) ripresa della parola da parte dell'educatore (McHoul, 1978; Mehan, 1979; Fele and Paoletti, 2003; Farini, 2008a). La frequenza con cui questa serie di azioni si verifica empiricamente segnala come i partecipanti all'interazione educativa si aspettino di essere coinvolti in una relazione asimmetrica tra un ruolo che accentra l'azione (l'educatore) e un ruolo limitato all'esperienza dell'azione altrui (gli educandi). Un'azione minimale degli educandi come l'annuire in risposta a un turno di parola dell'educatore, senza dire o fare altro, favorisce il prodursi di monologhi dell'educatore, svolgendo una funzione importante sia per l'interazione sia per l'analisi sociologica dell'interazione. Questa azione è però troppo veloce perché anche il più attento etnografo possa notarla e registrarla sul taccuino. Ancora più improbabile è che l'etnografo possa annotarla riproducendo con precisione la sua collocazione sequenziale che invece è molto importante, basti pensare alle conseguenze dell'annuire "fuori tempo" di uno studente, quando quest'azione è interpretata dall'insegnante come segnale di scarsa attenzione.
Oltre alla concreta possibilità di "perdere" azioni che rappresenterebbero tracce delle strutture sociali che orientano l'azione dei partecipanti all'interazione, un altro rischio connesso all'utilizzo delle note di campo è che il ricercatore interpreti una situazione in tempo reale come non farebbe se avesse la possibilità di osservarla altre volte, magari senza la pressione temporale che deriva dall'obbligo di registrare tutto mentre accade. L'assunzione di base della tecnica delle note di campo è che l'etnografo sia in grado catturare e rappresentare adeguatamente i più importanti fenomeni sociali; per fare questo è richiesta la capacità di determinare in tempo reale quelli che sono gli eventi importanti focalizzandovi l'attenzione a spese di altri che, non documentati da supporti audio-video, sono

destinati a scomparire definitivamente. Quello che l'etnografo armato di taccuino decide di non osservare non potrà più essere osservato.
Noi, però, consideriamo le interazioni che avvengono *realmente* e *naturalmente* nella società come oggetto di ricerca *sui generis*; in questa prospettiva sono necessarie tecniche di produzione dei dati più raffinate, in grado di preservare la complessità e la raffinatezza delle connessioni sequenziali tra azioni. Queste tecniche, che accompagnano l'AC fin dai suoi albori (Hutchby, 2007; Hutchby and Wooffitt, 2008; ten Have, 2008) consistono nella registrazione su supporto magnetico, o digitale, dell'interazione.

4.2 L'audio registrazione

L'audio registrazione dell'interazione fu la prima tecnica di produzione dei dati utilizzata dagli analisti della conversazione, ed è ancora oggi molto utilizzata.
L'audio registrazione offre notevoli vantaggi rispetto alle note di campo: alleggerisce il lavoro di registrazione dei dati richiesto al ricercatore e, preservando le interazioni su un supporto magnetico, o digitale, permette di osservarle ripetutamente. La possibilità di un'analisi ripetuta del dato grezzo crea le condizioni affinché siano notate cose che in precedenza erano sfuggite.
Rispetto alle note di campo l'audio-registrazione appare più intrusiva per la presenza dello strumento tecnico del registratore, che inoltre potrebbe essere soggetto a problemi di malfunzionamento. Tuttavia, sono oggi disponibili registratori di dimensioni ridotte e molto resistenti; allo stesso tempo, la tecnologia dell'audio-registrazione è ormai banale, e questo comporta un notevole abbassamento dei costi e la riduzione del rischio di difetti anche per i registratori più economici.
I problemi della tecnica dell'audio-registrazione, quindi, non consistono nella sua intrusività e nemmeno nei rischi di guasti o difetti dell'apparecchio di registrazione. I problemi non sono di ordine tecnico ma metodologico. In primo luogo, l'audio-registrazione non è in grado di produrre dati di qualità accettabile quando utilizzata per registrare interazioni dove il numero dei partecipanti è elevato, perché diventa impossibile identificare con sicurezza "chi fa cosa" con il solo supporto audio. Riprendendo l'esempio dell'interazione educativa, l'audio-registrazione non appare la tecnica più adatta a preservare la

complessità di interazioni che avvengono in aule con quindici, venti o anche trenta possibili parlanti.
Per ovvie ragioni di spazio, qui presentiamo esclusivamente i limiti dell'audio-registrazione quando essa è una tecnica praticabile. Pertanto, accenniamo solo al fatto che nessun registratore audio può registrare gli aspetti non verbali della comunicazione e che, quindi, nessuna ricerca interessata alla *multimodalità* dell'interazione (Goodwin, 2003; Stivers and Sidnell, 2005; Mondada, 2007; Monzoni and Drew, 2009) utilizzerà la tecnica dell'audio registrazione.
Prima di discutere della VO è importante richiamare due vantaggi che l'audio-registrazione offre rispetto a questa tecnica. In primo luogo, l'audio-registrazione non soffre del limite dell'unidirezionalità dello sguardo del mezzo tecnico. L'occhio della videocamera non è capace di una visione a 360 gradi che, invece, è alla portata di un buon microfono. In secondo luogo, le difficoltà tecniche connesse all'uso dell'audio-registrazione, seppur più rilevanti rispetto a quelle concernenti l'uso delle note di campo, non raggiungono quelle della VO.

4.3 La video-osservazione

Tuttavia, le riserve verso l'adozione della VO come tecnica di produzione-dati non sono di natura tecnica ma epistemologica, richiamandosi all'intrusività della videocamera che influenzerebbe il comportamento dei partecipanti all'interazione producendo, quindi, dati non attendibili.
Questa critica ci appare più legata al senso comune che non a concrete esperienze di ricerca sociale. I resoconti di numerose ricerche che adottano la VO concordano nell'affermare come, dopo un breve periodo, la presenza della videocamera non è più oggetto di attenzione (Baraldi, 2005; Iervese, 2006; Farini, 2008b; Fele, 2007). Inoltre, la videocamera è ormai un oggetto di uso comune e molte persone sono abituate a essere il soggetto di riprese video.
C'è un'altra osservazione, di carattere teorico, che ci induce a sminuire l'incidenza della videocamera sul comportamento dei partecipanti alle interazioni. Ammettendo l'esistenza di strutture in grado di orientare in modo relativamente stabile l'azione individuale (in caso contrario la sociologia non avrebbe senso!), queste strutture devono essere in grado di resistere alla perturbazione rappresentata

dalla presenza di una videocamera. Riteniamo molto improbabile che i partecipanti a un'interazione videoregistrata siano in grado di trasformarne "in tempo reale" i presupposti socioculturali e, anche se questo "miracolo sociologico" avvenisse, abbiamo ragioni per credere che lo farebbe in modo empiricamente osservabile, poiché la trasformazione dei presupposti socioculturali dell'interazione non può dipendere da un accordo telepatico tra i partecipanti. Essa deve essere oggetto di negoziazione, ovvero, nel caso sia iniziata unilateralmente da un partecipante, deve causare un visibile disorientamento, per breve che questo possa essere, nel suo interlocutore.
Inoltre, quale tipo di mutamento nel comportamento potrebbe generare la presenza della videocamera? Ci sono ragioni per credere come un eventuale mutamento del comportamento alla presenza di una videocamera tenderebbe all'espressione estremizzata di quelli che i soggetti considerano gli attributi del proprio ruolo. Queste ragioni derivano sia da esperienze empiriche di ricerca (Iervese, 2006) sia dall'idea che la presenza di un pubblico (la videocamera) stimoli un'"ansia da performance" che favorisce un più netto stagliarsi dei presupposti socioculturali che orientano l'azione dei partecipanti (valori fondamentali, ruoli diversificati, aspettative verso i risultati della comunicazione).
Ovviamente, poiché registrano la comunicazione nel momento in cui si produce, le VO sono utilizzabili fino al punto in cui non ostacolano la comunicazione. Si tratta non tanto del problema dell'intrusività del mezzo tecnico, ma del problema della disponibilità dei sistemi di comunicazione all'osservazione esterna, che riguarda tutte le tecniche di produzione dei dati. Come ogni tecnica di ricerca, anche la VO va utilizzata una volta valutati vantaggi e limiti. Dopo averne discusso i possibili rischi, vogliamo adesso esplorarne i vantaggi.

5 I dati della video-osservazione

La VO permette di cogliere, attraverso l'analisi dell'ordine sequenziale dell'interazione, come i partecipanti reagiscono alla condotta altrui e come reagiscono alle reazioni degli altri alle loro (Iervese, 2006). La VO è l'unica tecnica in grado di preservare su supporto magnetico o digitale lo sviluppo dell'interazione in tempo reale, rendendo così possibile osservarne i presupposti sociali e culturali nel momento in cui si concretizzano in specifiche azioni e

sequenze di azioni. Offrendo la possibilità di cogliere i più minuti aspetti dell'interazione, la VO è in grado di rendere giustizia alla complessità dei processi comunicativi (Goodwin, 2003).
Interrogativi importanti riguardano la natura dei dati che si producono adottando la tecnica della VO. Quali sono le loro caratteristiche? Sono dati appropriati per un'analisi dell'interazione attraverso l'AC?
La prima caratteristica dei dati che si producono mediante la tecnica della VO è che essi sono *audio* e *visivi*. Tutto quello che accade all'interno del raggio d'azione della videocamera è registrato acusticamente e otticamente; attraverso la VO è possibile tenere in considerazione tutto ciò che connota e contribuisce a dare forma all'interazione. Ovviamente, non tutto quello che accade in un contesto di interazione può essere registrato dalla videocamera. Abbiamo già accennato a come l'occhio della videocamera non sia capace di una visione a 360 gradi (4.2), e appare persino superfluo ricordare come anche il più raffinato dei microfoni può non cogliere i suoni più deboli. Possiamo però affermare, in accordo con Sacks (1992), che se la VO non garantisce una registrazione perfetta dell'interazione, essa garantisce una registrazione "abbastanza buona" di quel che è accaduto: di sicuro sono successe altre cose, ma perlomeno quello che è sul nastro è successo.
La VO produce sequenze di immagini dotate di fluidità che permettono di tenere conto dello svolgersi degli eventi e del loro concatenarsi in sequenze. La disponibilità di interazioni videoregistrate abilita l'analisi sociologica a riconoscere empiricamente la presenza di efficaci strutture dell'azione sociale sulla base del ripetersi di specifiche connessioni sequenziali tra azioni.
Un'altra caratteristica dei dati prodotti attraverso la VO è che essi, poiché sono fissati su supporti magnetici o digitali, sono disponibili per ulteriori osservazioni analitiche, anche molto tempo dopo la loro produzione. Ogni volta che si affronta nuovamente la sequenza videoregistrata, aspetti che prima erano stati trascurati posso attirare l'attenzione. Questa possibilità di ritornare ripetutamente sul dato originario permette una continua attività di affinamento dell'analisi e l'applicazione di una pluralità di piani di lettura (Zuengler, Ford and Fassnacht, 1998).
La memorizzazione dei dati ha un'altra conseguenza molto rilevante per la ricerca scientifica: svincola i dati dall'ipoteca della selettività del ricercatore che li raccoglie. Le note di campo riproducono solo

quello che l'etnografo ritiene importante nel momento in cui le redige; la videocamera, invece, registra tutto quello che è alla portata del suo obiettivo e del suo microfono.
La fissazione dei dati su supporti magnetici o digitali, infine, incide positivamente sulla socializzazione dei dati scientifici. Mentre le note di campo affidano al loro redattore un'ipoteca sui dati etnografici, giacché le sue selezioni sono irrevocabili e finiscono per confondersi con la complessità dei fenomeni, la VO consente la costruzione di estese raccolte di dati (relativamente) grezzi, disponibili per un'ampia gamma di interessi analitici.

6 Dalla teoria alla pratica

Pur se questo contributo, di natura metodologica, non ha lo scopo di discutere risultati di ricerca, riteniamo importante presentare un breve estratto di interazione raccolto in forma videoregistrata. L'interazione è stata videoregistrata nel corso una ricerca sulla comunicazione educativa nelle scuole primarie e medie inferiori, realizzata nel 2005[1] (Iervese, 2006).
La presentazione di un estratto di interazione vuole collegare discussione metodologica e pratica di ricerca sociale. Ovviamente, le osservazioni che seguono valgono per un'analisi dell'interazione multimodale, che tiene conto di tutte le risorse comunicative disponibili per i partecipanti e che ritiene come movimenti delle mani, sguardi, postura, movimenti del capo e così via rappresentino tipologie di azioni *sui generis*, che possono svolgere specifiche funzioni nell'interazione.
L'estratto riguarda una fase di un'attività educativa, dove gruppi di 4-5 bambini sono chiamati a costruire una storia di fantasia, basandosi su alcune foto. Un gruppo ha scelto di produrre una storia per immagini con fotografie e disegni, senza nessun testo scritto. L'educatore (edu) nota che una bambina (f1) non concorda con questa scelta, preferendo scrivere un testo; a questo punto egli interviene promuovendo un confronto tra i bambini per evitare il rischio di

[1] Il materiale video citato in questo contributo coinvolge minori ed è protetto dalla corrente normativa sulla privacy. Chi fosse interessato a prenderne visione per documentati interessi scientifici può contattare Federico Farini, all'indirizzo email federico.farini@unimore.it

emarginazione di f1. L'interazione è stata trascritta utilizzando le convezioni sviluppate da Jefferson, integrate da quelle che Goodwin ha sviluppato per la trascrizione della direzione dello sguardo (cfr. Du Bois *et al*, 1993). Il trascritto 1 si basa su una registrazione audio e video, il trascritto 2 è redatto sulla base della sola registrazione audio.

TESTO	il carattere maiuscolo indica volume alto
Testo	la sottolineatura indica enfasi particolare
°testo°	parola pronunciata ad u volume molto basso
test-	troncatura di un suono
te::sto	prolungamento di un suono (tanto maggiore quanti più sono i ':')
Hh	risatina (tanto più lunga quanti più sono gli 'h')
HH	risata (tanto più lunga quanti più sono gli 'H')
tes(h)to	parola pronunciata in modo ridente
(.)	pausa molto breve
[testo] [testo]	sovrapposizione tra parlanti
((commento))	commento trascrittore
((partecipante)) .------,	sguardo, in direzione del partecipante specificato tra doppie parentesi. Il punto indica il momento in cui lo sguardo viene orientato verso ((partecipante)), la virgola il momento in cui lo sguardo viene distolto
??? e ((??commento??)	identificazione del parlante e/o del tipo di azione comunicativa non verbale impossibile per mancanza di una registrazione video dell'interazione

Figura 1: Convenzioni di trascrizione del parlato (Jefferson, 1978 ; Goodwin, 1981)

1. Scuola Elementare San Martino in Rio (Re), 0405, *video*

1 edu: voi avete deciso di fare un cartellone con le immagini (.) ma se io adesso
2 vi chiedessi di inventare una storia, sulla base di queste cosa qua, vi
3 andrebbe di fare?
4 f1: sì, sì
5 edu: a me interessa vedere come arrivate alla decisione -
6 f1: chi vota per la storia?
7 edu: a me interessa sapere come arrivate alla decisione
8 m1: noi facciamo solo le votazioni
9 m2: facciamo una leggenda!
10 m1: facciamo un accordo (.) chi vota per la storia?
11 ((solo f1 alza la mano))
12 m3: hh HHH HHHH
13 m1: chi vota "no"?
14 ((quattro mani alzate))

15 m3: HH HHHH HH [all(h)ora: -
16 f1: [però io voglio fare la storia
17 m1: e va bene, la facciamo (.) la facciamo
18 op: quindi facciamo la storia?
19 m2: ((f1)) .----------,
OH, 'SCOLTA-
20 edu: ((m2)) .----------,
vince la maggioranza, dicono
21 f2: in[fat-]
22 edu: [NO] no scusa un attimo
23 m2: °no no°
24 f2: sì, infatti, dicono così, "vince la maggioranza", e prima abbiamo
25 incollato le foto che a qualcuno non andava!

2. Scuola Elementare San Martino in Rio (Re), 0405, *audio*

1 edu: voi avete deciso di fare un cartellone con le immagini (.) ma se io adesso
2 vi chiedessi di inventare una storia, sulla base di queste cosa qua, vi
3 andrebbe di fare?
4 f1: sì, sì
5 edu: a me interessa vedere come arrivate alla decisione -
6 f1: chi vota per la storia?
7 edu: a me interessa sapere come arrivate alla decisione
8 m1: noi facciamo solo le votazioni
9 ??: facciamo una leggenda!
10 m1: facciamo un accordo (.) chi vota per la storia?
11 ???
12 ??: hh HHH HHHH
13 m1: chi vota "no"?
14 ???
15 ??: HH HHHH HH [all(h)ora: -
16 f1: [però io voglio fare la storia
17 m1: e va bene, la facciamo (.) la facciamo
18 edu: quindi facciamo la storia?
19 ??: OH, 'SCOLTA- ((??a chi??))
20 edu: vince la maggioranza, dicono ((??a chi??)
21 ??: in[fat-]
22 edu: [NO] no scusa un attimo ((??perché interrompe??)
23 ??: °no no°
24 ??: sì, infatti, dicono così, "vince la maggioranza", e prima abbiamo
25 incollato le foto che a qualcuno non andava!

Che cosa offre la VO a una ricerca sociologica che ha per oggetto interazioni reali e naturali? Il confronto tra due trascrizioni della stessa interazione, la prima effettuata sulla base di registrazioni audio *e* video (1), l'altra sulla base di registrazioni solo audio (2), rende evidenti alcuni vantaggi offerti dalla tecnica della VO.

In primo luogo, la VO permette di riconoscere con (relativa) sicurezza i parlanti anche in interazioni multi-parte, con frequenti sovrapposizioni, interruzioni, rumori di fondo. Nel trascritto 2, invece, quattro turni di parola su un totale di venti non sono assegnabili ad alcun partecipante con un margine di errore accettabile. Questo ha importanti conseguenze negative sulla possibilità di verificare se, come e in che misura forme di ordine strutturale si manifestano nella distribuzione delle possibilità di partecipazione all'interazione.

In secondo luogo, la VO rende disponibili al ricercatore alcune tipologie di azioni interattive come sguardi e movimenti del corpo. L'incapacità del trascritto 2 di riprodurre la varietà delle risorse interattive utilizzate dai partecipanti impedisce di analizzare sequenze dove azioni non verbali svolgono specifiche funzioni, sia a livello dei rapporti di adiacenza tra azioni sia nella distribuzione delle possibilità di partecipazione (*turn taking*; cfr. Sacks, Schegloff and Jefferson, 1974).

Si prenda, ad esempio, la serie di turni 10-15. Nel trascritto 1, essa non pone particolari problemi di interpretazione: si svolge una votazione per alzata di mani, f1 finisce in minoranza ed è derisa dai compagni. Nel trascritto 2, invece, senza il supporto del dato videoregistrato diventa impossibile osservare la votazione per alzata di mano e il suo esito: la serie di turni appare assolutamente incomprensibile.

Anche l'impossibilità di osservare la direzione dello sguardo ha rilevanti conseguenze per l'analisi dell'interazione. Se il dato videoregistrato rende possibile comprendere che m2 si rivolge con tono esasperato a f1 (cfr. trascritto 1), quello che la sola audio-registrazione offre al ricercatore è che un bambino, peraltro non identificabile, risponde aggressivamente a una domanda dell'educatore (cfr. trascritto 2). Una simile forma della coppia adiacente "domanda dell'educatore-risposta dell'educando" potrebbe indurre a riconoscere una crisi delle strutture sociali dell'educazione, quando invece non è assolutamente il caso.

Infine, si faccia attenzione alla serie di turni 20-24. Il trascritto 1 permette di riconoscere che edu ha selezionato m2 come prossimo parlante attraverso lo sguardo. Quando f2 prende la parola abusivamente, edu sanziona questa azione interrompendola bruscamente prima che lei possa produrre un contributo sensato. Solo quando m2 esprime l'intenzione di non usufruire della possibilità di parlare concessagli da edu, f2 può riprendere la parola, arrivando a completare il proprio interevento senza interruzioni. Grazie alla ricchezza del dato videoregistrato, in questa serie di turni è possibile riconoscere empiricamente una struttura dell'educazione, ossia il ruolo dell'educatore come unico distributore delle possibilità di partecipazione tra gli educandi.
Una simile analisi appare relativamente semplice solo perché il trascritto 1 riproduce anche la direzione dello sguardo e permette di individuare i parlanti; molto più arduo è produrre la stessa analisi sulla base del trascritto 2. Quando il dato disponibile consiste in una registrazione solo audio, l'importante funzione dello sguardo come forma di selezione del prossimo parlante non è riconoscibile.

7 Conclusioni

L'AC descrive il flusso di azioni sociali compiuto dai partecipanti in interazioni di qualsiasi genere. Per "azione sociale" s'intende ogni comportamento che condiziona le azioni successive in conformità a procedure sociali vigenti, cui i partecipanti mostrino di essere orientati.
L'analisi sociologica dell'interazione richiede l'utilizzo di registrazioni in tempo reale che, preservando le connessioni sequenziali tra azioni, permettono di verificare empiricamente l'esistenza di forme di ordine sociale. A questo proposito, abbiamo argomentato come la VO sia la tecnica più efficace per produrre dati scientifici che preservano la complessità delle interazioni reali e naturali, permettendo quindi di apprezzare l'enorme mole di "lavoro sociale" necessario per il buon esito di attività apparentemente banali.
La VO permette di individuare con relativa sicurezza "chi fa cosa, e quando" anche in interazioni molto estese. Molti o pochi che siano i partecipanti all'interazione, la videocamera fissa su nastro anche il non-parlato, le posture, i gesti, gli sguardi. La disponibilità del dato videoregistrato, inoltre, offre all'osservatore la possibilità di cogliere,

ogni volta in cui si affronta nuovamente la sequenza videoregistrata, aspetti che in precedenza erano stati trascurati. Come affermato in precedenza (5), la possibilità di ritornare ripetutamente sul dato videoregistrato consente una continua attività di affinamento dell'analisi.

Analogamente al microscopio in biologia, la VO rende visibile quello che prima non lo era, ossia la complessità dell'interazione sociale. Grazie al loro elevato livello di dettaglio, e alla possibilità di analizzarli ripetutamente, i dati prodotti attraverso la tecnica della VO agevolano l'AC nel mettere in evidenza quelle regolarità nei comportamenti umani che permettono di dedurre, su basi genuinamente empiriche, la presenza di efficaci strutture dell'azione sociale.

Riferimenti bibliografici

Atkinson, J.M. and J. Heritage (1984) *Structures of social action: studies in Conversation Analysis*. Cambridge: Cambridge University Press.

Baraldi, C. (2005) *Costruire la diversità e il dialogo con bambini e preadolescenti*. Imola: La Mandragora.

– (2008) *Dialogare in classe. La relazione tra insegnanti e studenti*. Roma: Donzelli

Baraldi, C. and L. Gavioli (2007) 'The relevance of interactions in functionally differentiated society: the contribution of conversation analysis to the theory of social systems'. *Soziale Systeme* 13(1-2), pp. 125-35.

Du Bois, J.W., S. Schuetze-Coburn, S. Cumming and D. Paolino (1993) Outline of discourse transcription. In J.A. Edwards and M.D. Lampert (eds) *Talking data: Transcription and coding in discourse research*, Hillsdale, NJ: Erlbaum, pp. 45-89.

Farini, F. (2008a) 'L'espressione della diversità dei punti di vista in classe: risorsa o pericolo per l'educazione?' *Orientamenti Pedagogici* 55(5), pp. 877-98.

– (2008b) 'Pratiche di promozione della partecipazione attiva: analisi di un intervento sociale rivolto ad adolescenti e giovani adulti nel Distretto 1 della Provincia di Modena'. *La rivista di servizio sociale* 1, pp. 34-59.

– (2009) ‘An improbable leadership: structural limits of educational communication’. *International Journal of Leadership in Education* 12,(4), pp. 367-88.
Fele, G. (2007) *L’ analisi della conversazione*. Bologna: il Mulino.
Fele, G. and I. Paoletti (2003) *L’interazione in classe.* Bologna: il Mulino.
Garfinkel, H. (1967) *Studies in ethnomethodology*. Englewood Cliffs, NJ: Prentice Hall.
Giglioli, P. and G. Fele (2000) *Linguaggio e contesto sociale*. Bologna: il Mulino.
Goodwin, C. (1981) *Conversational Organization: Interaction Between Speakers and Hearers*. New York: Academic Press.
– (2003) *Il Senso del Vedere: Pratiche Sociali della Significazione*. Roma: Meltemi Editore.
Hutchby, I. and R. Woffitt (2008) *Conversation Analysis*. London: Polity Press.
Hutchby, I. (2007) *The Discourse of Child Counselling*. Amsterdam: John Benjamins.
Iervese, V. (2006). *La gestione dialogica del conflitto*. Imola: La Mandragora.
Jefferson, G. (1978) Sequential aspects of storytelling in conversation. In J.N. Schenkein (ed.) *Studies in the organization of conversational interaction*, New York: Academic Press, pp. 213-48.
Luhmann. N. and R. De Giorgi (1992) *Teoria della società*. Milano: Franco Angeli.
Luhmann, N. (1984) *Soziale Systeme. Grundriß einer allgemeinen Theorie*. Frankfurt a. M.: Suhrkamp.
McHoul, A. (1978) ‘The organization of turns at formal talk in the classroom’. *Language in Society* 7, pp. 183-213.
Mehan, H. (1979) *Learning lessons: Social organization in the classroom*. Cambridge (MA): Harvard University Press.
Mondada, L. (2007) ‘Multimodal resources for turn-taking: Pointing and the emergence of possible next speakers’. *Discourse Studies* 9(2), pp.195-226.
Monzoni, C.M. and P. Drew (2009) ‘Inter-interactional contexts of story-interventions by non-knowledgeable story recipients in

(Italian) multi-person interaction'. *Journal of Pragmatics* 41, pp. 197-218.
Sacks, H. (1992) *Lectures on Conversation I & II* (ed. by G.Jefferson). Oxford: Blackwell.
Sacks, H., E.A. Schegloff and G. Jefferson (1974) 'A Simplest Systematics for the Organisation of Turn-Taking for Conversation'. *Language* 50, pp. 696-735.
Schegloff, E.A. and H. Sacks (1973) 'Opening Up Closings'. *Semiotica* 8, pp. 289-327.
Spencer-Brown, G. (1969) *Laws of Form*. London: Allen and Unwin.
Stivers, T. and J. Sidnell (2005) 'Introduction: Multi-Modal Interaction'. *Semiotica* 156(1/4), pp. 1-20.
ten Have, P. (2008) *Doing Conversation Analysis*. London: Sage.
Zuengler, J., C. Ford and C. Fassnacht (1998) *Analyst eyes and camera eyes.* New York: The National Research Center on English Learning and Achievement.

Sbagliando si impara
Strategie per l'attenuazione dell'errore in un sistema di riconoscimento *speaker independent* per la lingua italiana

Maria Palmerini
Cedat 85
m.palmerini@cedat85.com

Abstract

In this paper a research conducted in 2008 by Cedat 85 (leader firm in speech processing) in cooperation with EML (European Media Laboratory, in Heidelberg) will be presented. The project was aimed to create a speaker independent speech recognition system for Italian spontaneous speech. After a short description of the main features of the system and of the data used for the training, we would like to focus on the whole activity of error analysis and reduction. We will describe different error typologies found in a fist test; then, we will address to some of the possible operations that can be done on the system in order to reduce errors in transcriptions: as it will be shown, some of these operations can be done directly on the system, while some of them can be done in a second phase, on transcriptions.
One of the aspect that we would like to highlight is that, although error reduction is very important, mistakes will never be completely eliminated; besides, such a result wouldn't be really useful anyway, if we think of the main applications where a speech recognition system could be used, simply because a transcription containing each single element of speech is not what is really needed, most of the times.
Finally, we indicated some of the most recent improvements of the system and described the last research activities.
In the paper some results obtained in first experiments, and some images of the software used to correct and analyse errors will be presented. These results and some statistical information related to the language model will be resumed in tables in the Appendice, in the end of the paper.

Introduzione

La tecnologia del riconoscimento vocale (o *speech recognition* nella letteratura del settore) ha compiuto, negli ultimi anni, significativi passi in avanti dovuti sia al raffinamento delle tecnologie impiegate sia alla notevole potenza elaborativa sempre più disponibile anche nei semplici computer personali. Dai primi programmi creati per riconoscere solo singole parole su vocabolari limitati si è passati a sistemi in grado di trascrivere il parlato continuo di un determinato parlatore, con vocabolari molto più vasti, fino a raggiungere la

capacità di riconoscere il parlato spontaneo di un qualsiasi parlatore senza bisogno di addestramento su uno specifico profilo vocale.
In questo articolo presenteremo un progetto di ricerca condotto da Cedat 85, azienda leader nel settore del trattamento del parlato, in collaborazione col centro EML (European Media Laboratory) di Heidelberg, da aprile a dicembre 2008.
Grazie all'ampia banca di dati raccolta in molti anni di attività nel settore della resocontazione, Cedat 85 ha messo a punto un sistema di riconoscimento del parlato spontaneo indipendente dal parlatore (*spontaneous speech, speaker independent*) per la lingua italiana. Il sistema si basa sulla tecnologia recentemente realizzata da IBM Research (Yorktown) e denominata VoiceTailor. Il sistema è stato accuratamente testato e dall'individuazione dei primi errori si è proceduto a una loro classificazione per tipologie; successivamente, sono state progettate e in parte realizzate strategie di riduzione dell'errore per ottenere un risultato migliore nella trascrizione.
Presenteremo in questa sede i primi risultati, sebbene il progetto sia ancora in corso.
Per una descrizione delle tecnologie di ultima generazione impiegate nel sistema di riconoscimento VoiceTailor, rimandiamo alla disponibilità di documentazione come pubblicata da IBM.
Gran parte della terminologia che utilizzeremo è mutuata dall'inglese, pertanto saranno inserite di volta in volta in nota eventuali chiarimenti.

1 Il sistema di riconoscimento vocale di Cedat 85

Come nella gran parte dei sistemi di riconoscimento vocale, inclusi quelli di ultima generazione, la tecnologia impiegata nel sistema messo a punto da Cedat 85 per la lingua Italiana utilizza modelli di previsione di tipo markoviano (HMM) e, in generale, si basa prevalentemente su criteri statistici.
Semplificando e per meglio analizzare le parti che seguono, il sistema è composto da un modello acustico, deputato al riconoscimento dei foni, e da un modello di linguaggio, grazie al quale vengono identificate le parole singole e in sequenza. Data la natura statistica di base, un ruolo fondamentale, per il buon funzionamento del sistema, è giocato dai dati di addestramento di questi due modelli, che vedremo più nel dettaglio.

1. 1 Il modello acustico

Il modello acustico[1] è stato costruito su 150 ore circa di parlato campionato a 16KHz. Si tratta di registrazioni di assemblee istituzionali di vario genere: in gran parte convegni e assemblee politiche, ma anche alcune trasmissioni televisive di attualità (prevalentemente telegiornali).
Il parlato contenuto in queste registrazioni è in gran parte spontaneo o parzialmente pianificato all'interno di eventi comunicativi piuttosto ordinati; sono comunque state eliminate tutte le sovrapposizioni di più voci, le scarse parti che contenevano dialetto, le voci non native. Il corpus acustico contiene circa 940[2] voci di parlanti diversi, tutti adulti e madrelingua italiani, dei quali circa il 25% sono voci femminili. Gli accenti regionali presenti coprono praticamente tutto il territorio nazionale.
Tutto questo materiale audio è stato trascritto manualmente e le trascrizioni sono state allineate all'audio ogni 15-20 secondi.

1.2 Il modello di linguaggio

Il modello di linguaggio è stato creato con l'obiettivo di riconoscere il linguaggio utilizzato nelle assemblee di tipo politico istituzionale. Dopo l'analisi di un consistente corpus relativo all'ambiente di addestramento, si è selezionato un vocabolario di circa 30.000 entrate[3]; le sequenze considerate per la costruzione delle statistiche sono composte da una, due e tre parole consecutive. Nella

1 I dati riportati sono quelli di addestramento del modello utilizzato nella fase di sperimentazione, al quale si riferiscono i dati sugli errori riportati in questo articolo. Attualmente, il modello acustico è stato ulteriormente ampliato; si rimanda, per i dettagli, ai paragrafi successivi.
2 Sul numero dei parlanti non possiamo fornire dati precisi poiché alcuni di essi sono rimasti non identificati.
3 Col termine "entrata" del vocabolario si intende ogni parola graficamente diversa da un'altra, indipendentemente dal lemma a cui appartiene o dalla sua funzione sintattica. Per fare un esempio, il testo "Gianni lava la mela, la sbuccia e la mangia" contiene 7 entrate di vocabolario, dove le tre occorrenze di "la" contano come una sola entrata.

terminologia del settore, dunque, può essere definito un modello di linguaggio a trigrammi[4].
Il materiale utilizzato per la creazione del modello di linguaggio è costituito da testi scritti per 220 milioni di parole[5] circa. Il corpus comprende: tutti i resoconti di assemblee istituzionali e non, redatti da Cedat 85 nel corso della sua attività negli anni 2003-2008; resoconti dello stesso genere scaricati da internet; articoli di attualità e politica da quotidiani; le trascrizioni di tutto l'audio contenuto nel modello acustico.
È evidente, dunque, che si tratta di un modello di linguaggio fortemente centrato sul linguaggio assembleare di tipo politico; come vedremo, questo aspetto ha conseguenze decisive sul riconoscimento che darà i risultati migliori proprio su parlato di questa tipologia.

1.3 Il test set

Dall'insieme dei dati disponibili per la messa a punto del sistema, abbiamo isolato un campione che non è stato utilizzato per l'addestramento, ma solo per testare il riconoscimento.
Questo piccolo corpus, che chiameremo *test set*, è costituito da circa 6 ore di parlato con le stesse caratteristiche acustiche del materiale utilizzato per l'addestramento, appartenente a 21 parlanti diversi fra loro e diversi dalle voci utilizzate per l'addestramento; di questi, 17 sono uomini e 4 sono donne.
Il sistema di trascrizione ha avuto una sua dinamica realizzativa, dove ai primi modelli ne sono succeduti altri con la messa a punto di parametri e di strutture addestrative; questo meccanismo è continuo e in parte inserito direttamente nel sistema. Chiaramente le

4 Questo termine, che etimologicamente potrebbe far pensare a una sequenza di tre caratteri grafici, invece nella terminologia del settore indica una sequenza di tre parole; come altre, è mutuato dalla terminologia anglosassone e in particolare dai termini *unigram*, *bigram*, *trigram*, *n-gram* che quindi in italiano vengono resi coi rispettivi unigramma, bigramma, trigramma, n-gramma.
5 Per "parola" si intende un'unità grafica conposta da caratteri alfabetici che va da spazio a spazio che viene contata in ogni sua occorrenza, indipendentemente dal lemma a cui appartiene o dalla sua funzione sintattica. Per utilizzare l'esempio già citato, il testo "Gianni lava la mela, la sbuccia e la mangia" è composto da 9 parole, dove le tre occorrenze di "la", che pure ha due funzioni morfosintattiche diverse, contano come tre parole.

considerazioni sugli errori che seguono fanno riferimento ad una precisa fase nella evoluzione del sistema, sistema attualmente piuttosto diverso e migliorato in molti aspetti.

2 Primi risultati: individuazione degli errori

Un primo test, dunque, è stato fatto sottoponendo al sistema il *test set* sopra descritto; dalla trascrizione prodotta dal programma di riconoscimento, sono stati estratti e analizzati gli errori, che sono stati, innanzitutto, suddivisi in tre tipologie. Ne daremo una descrizione nei paragrafi seguenti.
Un sommario sui risultati del test viene riportato in una apposita tabella in appendice.

2.1 Sostituzioni

Parliamo di sostituzioni quando una parte del parlato è stata decodificata erroneamente e dunque nello scritto è stata sostituita da una o più parole non corrispondenti. Sul totale degli errori, le sostituzioni sono il numero maggiore (56,9%); nel nostro test, sul totale delle parole trascritte rappresentano il 9%.
Vediamone alcuni esempi.

a)
Pronunciato: sino ad oggi RETTO fondamentalmente solo dagli accordi della organizzazione
Trascritto: sino ad oggi ELETTO fondamentalmente solo dagli accordi della organizzazione

b)
Pronunciato: chiediamo che sia armonizzato con LE DEFINIZIONI del regolamento regionale
Trascritto: chiediamo che sia armonizzato con LA DEFINIZIONE del regolamento regionale

c)
Pronunciato: rettore dell' università di COSTANTIN in ALGERIA
Trascritto: rettore dell' università di COSTRETTI in PERIFERIA

Come si vede, le sosituzioni possono investire una singola parola (esempio a), ma anche una sequenza di parole (esempi b e c). In particolare, l'esempio citato in (c) mostra uno degli aspetti caratteristici dei sistemi di riconoscimento vocale che utilizzano criteri statistici per le previsioni, ovvero l'"esplosività dell'errore". In altre parole, poiché le previsioni partono dall'osservazione del contesto (sia pure limitato a sequenze di tre parole), ogni errore altera il contesto nel quale la parola è inserita e può, dunque, condizionare negativamente le aspettative sulle parole adiacenti, causando nuovi errori.
Nella maggior parte dei casi, però, come si osserva negli esempi, le sostituzioni derivano dalla difficile decodifica delle vocali iniziali e finali che, notoriamente, nel parlato spontaneo sono ipoarticolate o comunque sono investite da fenomeni di coarticolazione.

2.2 Cancellazioni

Una seconda tipologia di errori è data dalle parole che, pur essendo pronunciate dall'oratore, non vengono individuate dal sistema e quindi non compaiono nella trascrizione.
Nel nostro test il 3,6% delle parole pronunciate sono state cancellate; queste rappresentano il 21,9% degli errori.
Si riportano sotto alcuni esempi.

a)
Pronunciato: HA attecchito stentatamente...
Trascritto: *** attecchito stentatamente...

b)
Pronunciato: veniamo all' argomento DEL nostro convegno...
Trascritto: veniamo all' argomento *** nostro convegno...

c)
Pronunciato: CARATTERIZZATI NON soltanto DALL' ASSENZA DI libertà ma anche...
Trascritto: CARATTERIZZATO IN soltanto *** SENZA *** libertà ma anche...

Nel primo esempio, è ragionevole supporre che il verbo "ha" non sia

stato davvero articolato e quindi non sia affatto presente nel segnale se non – nella migliore delle ipotesi – sotto forma di un leggero allungamento della vocale iniziale di "attecchito". Nell'esempio (b) il segmento cancellato ha una maggiore consistenza fonetica, ma può comunque essere stato ridotto a causa di fenomeni coarticolatori.
Nell'ultimo esempio, invece, i segmenti cancellati sono due, pertanto la dimensione del fenomeno è maggiore. Anche in questo caso, si osserva che nel segmento interessato (la seconda parte della frase) si sono accumulati più errori, in particolare abbiamo una cancellazione seguita da una sostituzione, seguite da una seconda cancellazione. È difficile dire con certezza quale di questi errori abbia dato inizio ai successivi; il dato certo è, comunque, che in questi casi errori adiacenti sono strettamente legati uno all'altro. Ad esempio, è possibile che l'erronea trascrizione di "senza" al posto di "assenza" abbia imposto la cancellazione delle due preposizioni, che altrimenti avrebbero generato un trigramma non accettabile ("dalla senza di").

2.3 Inserzioni

Si parla di inserzione quando il sistema trascrive con una parola un segmento di audio dove non è rilevabile la pronuncia di quella parola. Solitamente, errori di questo tipo sono generati dalla presenza di esitazioni non lessicali (tipicamente un "mmmh" genera l'inserzione di un "non" in trascrizione), rumori o musica di fondo.
Nel nostro test, il 3,4% delle parole trascritte non ha un corrispettivo nell'audio, dunque sono state inserite dal sistema. Sul totale degli errori presenti, il 21,1% è costituito da inserzioni.
Anche in questo caso, riportiamo alcuni esempi.

a)
Pronunciato: il tema di fondo è *** se quella idea dei diritti dei doveri sia rimasta in campo...
Trascritto: il tema di fondo è CHE se quella idea dei diritti dei doveri sia rimasta in campo...

b)
Pronunciato: e di un certo civismo che dovrebbe accompagnare *** un paese
Trascritto: e di un certo civismo che dovrebbe accompagnare PER un

paese

c)
Pronunciato: visto che *** vi state occupando di questi SESSANT' anni...
Trascritto: visto che QUI vi state occupando di questi SESSANTA anni...

In tutti questi esempi, è interessante osservare i segmenti inseriti – che comunque solitamente sono singole parole – vanno ad aggiungersi in sequenze che comunque, limitatamente al dominio delle tre parole consecutive[6] – risultano accettabili. Per questo motivo, il sistema ha potuto assegnare loro un punteggio alto, in termini di probabilità.

3 Principali cause individuate

Dopo aver classificato le diverse tipologie errore, abbiamo cercato di capire quali fossero le principale cause.
Da un ascolto dell'audio utilizzato come *test set*, è emerso che una buona parte degli errori di trascrizione sono dovuti a una cattiva pronuncia dello speaker; ogni esitazione, autointerruzione, autocorrezione, parola mal pronunciata porta inequivocabilmente con sé almeno un errore, come si vede nell'esempio seguente.

a)
Pronunciato: perché non è possibile che UNA FRESTRUTTURA che debba essere assolutamente neutrale
Trascritto: perché non è possibile che UN' ALTRA STRUTTURA che debba essere assolutamente neutrale

Se analizziamo gli errori secondo le tipologie che abbiamo descritto sopra, per quanto riguarda inserzioni e cancellazioni, possiamo dire che in effetti il sistema non si comporta molto diversamente dalla facoltà umana di comprensione del linguaggio: poiché il parlato spontaneo è ipoarticolato, il sistema – così come la parte del nostro cervello deputata alla comprensione linguistica – non ricorre solo alle

6 Ad esempio, nella frase riportata in (c), i trigrammi "visto che qui", "che qui vi" e "qui vi state" sono tutti perfettamente accettabili in italiano.

informazioni contenute nel segnale, ma le integra con altre informazioni che in un sistema automatico sono principalmente di tipo statistico, mentre nel nostro cervello sono anche di tipo sintattico, semantico e pragmatico.
Pertanto, il sistema 'impara' che, anche laddove non vi siano informazioni nel segnale, a volte occorre inserire parole solo perché queste possono essere recuperate dal modello di linguaggio (come nel caso, poniamo, di "è evidente", in cui ci aspettiamo che il verbo "è" venga comunque trascritto anche se non fosse affatto articolato). Gli errori di inserzione, dunque, spesso possono essere considerati eccessi di ricostruzione linguistica per far fronte alla scarsità di informazioni nel segnale.
Similmente, le cancellazioni possono essere dovute a un eccesso nell'ignorare informazioni sì presenti nel segnale, ma solitamente non significative (rumori di sottofondo, false partenze, musica, esitazioni non lessicali). Per questo, così come il sistema 'impara' che non deve scrivere "ma" o "no" ogni volta che sente una consonante nasale non lessicale (ad esempio "mmmh"), per la stessa ragione potrà cancellare un "non" assolutamente indispensabile alla comprensione ma ipoarticolato.
Passando alle sostituzioni, la principale causa è data dalle parole che non sono contenute nel vocabolario. Come è ovvio, infatti, il sistema non è in grado di generare parole autonomamente, ma si limita, dato un certo input acustico, a cercare la stringa che più gli assomiglia tra quelle che ha nel suo *data-base*. Se uno speaker pronuncia una parola sconosciuta al sistema, non vi è alcuna possibilità che questo la trascriva correttamente. Questa la ragione dell'errore contenuto nell'esempio seguente:

b)
Pronunciato: wireless smart phone open source questi alcuni degli imperativi HIGH TECH del due mila nove
Trascritto: wireless smart phone open source questi alcuni degli imperativi AI TESTI del due mila nove

Come si è visto (sempre semplificando), la produzione del testo trascritto è il risultato di un processo di decodifica del segnale acustico durante il quale vengono utilizzati i due modelli che costituiscono la

conoscenza linguistica del sistema: il modello acustico e il modello di linguaggio. Questi interagiscono uno con l'altro e contribuiscono in ugual misura alla creazione delle statistiche; tuttavia, alcuni errori mostrano che in certi casi uno dei due modelli può prevalere sull'altro. Vediamo i due esempi che seguono:

c)
Pronunciato: principalmente per la WEB NAVIGAZIONE
Trascritto: principalmente per la QUOTA DI AZIONI

d)
Pronunciato: ma non L' hanno visto in televisione L' hanno visto sul loro COMPUTER
Trascritto: ma non ** hanno visto in televisione ** hanno visto sul loro COMPIUTA

Rispetto al primo esempio, segnaliamo che sia l'entrata "web" sia l'entrata "navigazione" erano contenute nel vocabolario ed erano correttamente descritte fonologicamente; il modello acustico, dunque, avrebbe dovuto essere in grado di riconoscerle. Tuttavia, in un modello di linguaggio italiano, la sequenza "web navigazione" non è comune, poiché ricalca una struttura sintattica di tipo anglosassone e, quindi, una sequenza statisticamente poco probabile in italiano. È possibile dunque che, secondo le informazioni possedute dal modello di linguaggio, la stringa non sia stata ritenuta statisticamente accettabile. Il risultato è una trascrizione che riporta una sequenza di parole plausibile, ma molto meno pertinente dal punto di vista acustico.
Nell'esempio in (d), invece, sembra essersi verificato il fenomeno contrario. Tra le due ipotesi – "computer" vs "compiuta" – il sistema ha scelto quella più vicina[7] all'input acustico, anche a costo di generare una stringa ("sul loro compiuta") non plausibile statisticamente per il modello di linguaggio.

7 Occorre precisare che al momento del test l'entrata 'computer' era stata associata a una trascrizione fonologica automatica secondo le regole fonologiche della lingua italiana, ovvero [komputer]; in seguito a questo errore, la descrizione fonologica è, dunque, stata corretta manualmente in [kompjuter].

4 Strategie di riduzione dell'errore

L'individuazione degli errori e l'analisi delle cause ha lo scopo finale di fornire riflessioni e informazioni utili per il miglioramento del sistema.
Poiché le cause sono non sempre note e piuttosto differenziate, anche gli interventi finalizzati alla riduzione dell'errore dovranno investire più fronti, con la consapevolezza – da parte nostra e dell'utente finale – che l'errore non potrà comunque essere eliminato del tutto.

4.1 Interventi sui modelli e sul vocabolario

Innanzitutto, il test ha rivelato la necessità di rafforzare i principali componenti del sistema di trascrizione, ovvero i due modelli (acustico e di linguaggio) e il vocabolario.
Per quanto riguarda il modello acustico, è possibile operare una serie di interventi cosiddetti "di *tuning*" su alcuni parametri. Su questo aspetto, molto complesso, non ci addentreremo in questa sede, mentre ci concentreremo più sugli interventi di tipo linguistico che nel modello acustico riguardano la fonologia: saranno quindi perfezionate la rappresentazione fonologica delle singole entrate e la descrizione dei fonemi della lingua italiana, così da agevolare la corretta associazione tra un certo input acustico e la relativa trascrizione. Va ricordato che la fonologia di un sistema di trascrizione automatica si basa sui dati forniti dalla letteratura specializzata nel settore, ma può talvolta discostarsi da questa; la ragione è che le informazioni sulla fonologia devono essere espresse in una forma tale da renderli utilizzabili da parte degli algoritmi del sistema, così da consentire a quest'ultimo di gestire, ad esempio, la variabilità di realizzazione fonetica di una stessa parola fra un parlante e un altro.
Un primo intervento necessario a migliorare il riconoscimento del parlato è sicuramente l'ampliamento del vocabolario grazie all'inserimento di nuove entrate. Non vi sono limiti apparenti alle dimensioni di un vocabolario, per la lingua italiana ve ne sono di oltre 200.000 entrate. Su questo punto, però, è opportuno fare alcune considerazioni legate sia alle tecnologie utilizzate sia agli obiettivi che intende raggiungere chi costruisce l'applicazione.
È evidente che quante più sono le parole conosciute dal sistema, tanto più complessa risulterà sia la selezione dei candidati per ogni istanza

acustica, sia l'analisi contestuale delle possibili sequenze di parole. È altrettanto evidente che tante più sono le entrate inserite nel vocabolario, tanto più grande dovrà essere il corpus di addestramento per poter avere evidenze statistiche significative nella costruzione del modello di linguaggio. Mentre il primo problema è principalmente relativo alla potenza di elaborazione disponibile e alla possibile parallelizzazione degli algoritmi impiegati, il secondo è legato alla disponibilità di testi scritti nel settore di riferimento (dobbiamo considerare adeguati corpora per miliardi di parole quando si superano le 100.000 entrate nel vocabolario). In entrambi i casi vi sono dei costi che possono non compensare il beneficio di un vocabolario sovradimensionato.

La scelta della dimensione di un vocabolario è un processo non semplice e che richiede competenze ed esperienze specifiche. Selezionando un settore di riferimento (ad esempio politica, sport, cronaca, medicina, informatica ...) e analizzando il materiale scritto relativo si vede subito che il numero di parole diverse impiegato non è particolarmente ampio. Teniamo presente che l'aumento delle parole diverse, invece, è molto spesso legato a nomi, cognomi, luoghi geografici, termini tecnici e consimili, entità che non sono limitate nel numero, ma che molto spesso hanno un utilizzo basso e limitato ad archi temporali piuttosto brevi[8]. Quindi la scelta delle entrate da inserire nel vocabolario non è mai semplice e dipende da molti fattori quali, ad esempio, il settore di riferimento dell'applicazione, una analisi approfondita dei testi disponibili relativi al settore, considerazioni sulle frequenze di occorrenza delle parole, tipologia delle stesse, considerazioni sui risultati attesi dall'applicazione e analisi delle risorse richieste.

A proposito dei modelli acustico e di linguaggio, poi, è opportuno precisare che il sistema di trascrizione permette di selezionare un determinato modello fra tutti quelli disponibili per ogni richiesta di trascrizione che gli viene sottoposta, quindi non è assolutamente necessario avere modelli universali ma è molto più efficace avere più modelli specifici per i settori che via via sono utilizzati.

8 Nel linguaggio televisivo, ad esempio, ma anche in quello politico, possiamo citare il caso del nome proprio "Eluana" che ha avuto per un certo periodo frequenza altissima, ma che al di fuori di quella precisa fascia temporale non è affatto un nome frequente in italiano.

Tutto questo porta a considerare l'attività di aggiunta di entrate al vocabolario, dopo la rilevazione di errori, non semplice e lineare date le implicazioni che porta con sé.
In ogni caso, l'ampliamento del vocabolario non è sufficiente per migliorare il riconoscimento; come si è visto nell'esempio citato in (c) nel paragrafo 3, il fatto che sia l'entrata "web" sia l'entrata "navigazioni" fossero presenti nel vocabolario non ha impedito l'errore, perché le due parole non occorrevano mai in quella sequenza nel modello di linguaggio; pertanto, il bigramma "web navigazioni" risultava sconosciuto al sistema. In alcuni casi, dunque, è utile curarsi che siano rappresentate non solo le singole parole, ma anche i bigrammi e trigrammi possibili.

4.2 Collaborare con l'utente

Ogni prodotto tecnologico, dopo essere stato realizzato, richiede una fase di sperimentazione indispensabile per individuare eventuali difficoltà nell'utilizzo pratico reale. Una delle idee più innovative del progetto realizzato da Cedat 85 è sata proprio quella di cercare di collaborare con gli utenti per raccogliere informazioni preziose per impostare l'attività di miglioramento del sistema.
È stato dunque ideato e sviluppato un software con funzioni di *client* (figura 1) grazie al quale l'utilizzatore può inviare un file audio al sistema, riceverne la trascrizione e correggere gli errori in un ambiente apposito. Grazie a questo software, inoltre, ogni correzione o modifica apportata dal cliente al file di testo viene salvata su un file xml che poi viene raccolto dagli addetti alla manutenzione del sistema.

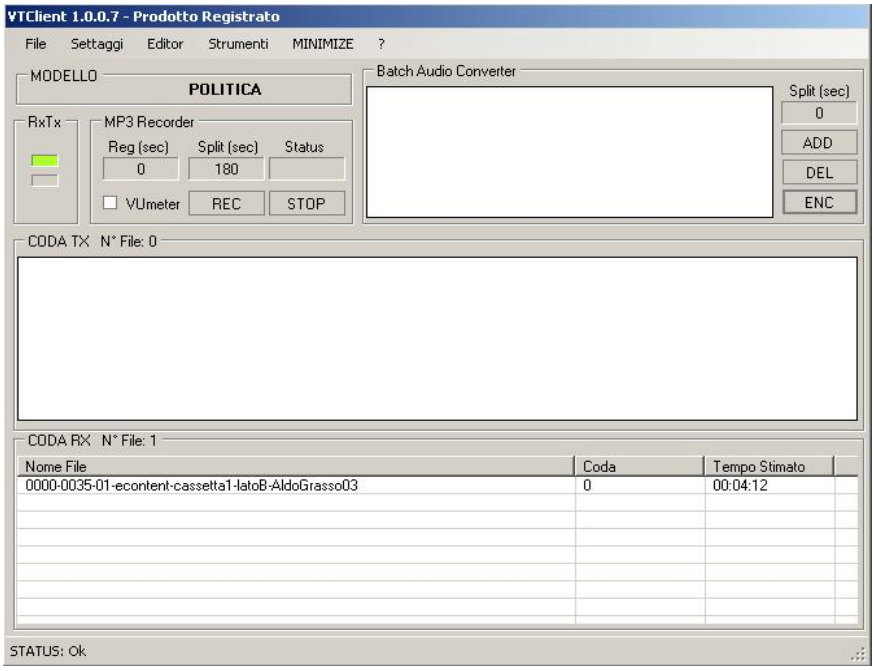


Figura 1: Il *client* sviluppato da Cedat 85 per l'invio dei file audio e la ricezione della trascrizione automatica.

In questo modo, ogni correzione apportata dall'utente (figura 2) può essere visionata dagli operatori di Cedat 85, i quali, grazie a un altro software sviluppato internamente all'azienda e in grado di processare i file xml e di estrarne i dati relativi alle modifiche del testo, hanno tutte le informazioni necessarie per poter operare alcuni interventi di miglioramento del sistema di trascrizione.

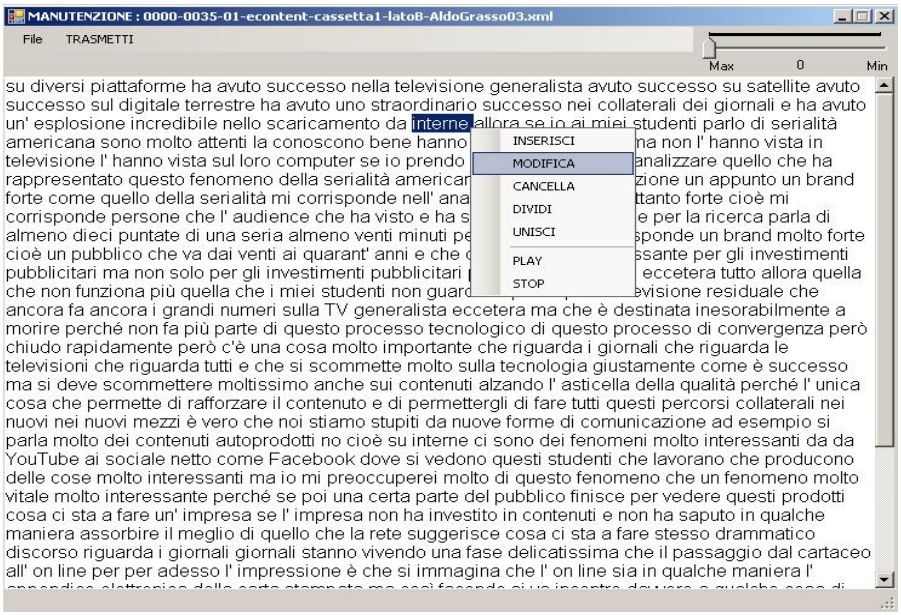


Figura 2. Il *client* sviluppato da Cedat 85 nell'area di editing.

È possibile analizzare automaticamente le modifiche apportate a uno o più documenti ed avere le statistiche ad essi relative; in particolare è utile conoscere: le parole non presenti fra le entrate del vocabolario; l'elenco dei vari tipi di errore suddivisi per tipologia, con la possibilità di vederne qualità e quantità; è inoltre sempre possibile riascoltare l'audio originale nel punto in cui gli errori si sono verificati.
Grazie alla messa a punto di questi strumenti, è stato possibile realizzare un secondo test. Sono stati inviati al sistema di trascrizione alcuni file audio della durata complessiva di circa 6 ore, che contenevano circa 42.000 parole. Nella tabella 1 sono rappresentati i primi risultati delle statistiche ottenute.

SOSTITUZIONI	CANCELLAZIONI	INSERZIONI
72 -> è ==> e	181 -> e	183 -> e
35 -> e ==> è	127 -> è	96 -> per
21 -> d' ==> di	59 -> che	96 -> è
21 -> e ==> che	56 -> a	61 -> il
16 -> la ==> della	48 -> l'	50 -> di
15 -> che ==> e	47 -> di	49 -> che
14 -> del ==> della	47 -> il	47 -> a
12 -> il ==> del	39 -> in	43 -> un
11 -> alla ==> la	30 -> ha	41 -> la
11 -> antitrust ==> trust	28 -> ma	31 -> o
11 -> ha ==> a	27 -> un	28 -> in
10 -> all' ==> l'	26 -> o	25 -> al
10 -> cioè ==> c'è	24 -> la	25 -> non
10 -> della ==> la	21 -> non	22 -> del
10 -> l' ==> dell'	20 -> si	22 -> l'
10 -> la ==> alla	19 -> i	20 -> le
10 -> nel ==> del	15 -> per	17 -> i
	13 -> al	16 -> lo
	12 -> del	15 -> da
	12 -> se	14 -> ha

Tabella 1. Dettagli sugli errori più frequenti.

Come si vede dai dati riportati, gli errori più frequenti[9] riguardano soprattutto parole grammaticali di poca consistenza fonologica e, quindi, poco riconoscibili dal punto di vista acustico. Per ridurre questo genere di errori si è scelto, dunque, di intervenire sia rafforzando il modello di linguaggio, sia ricorrendo all'utilizzo di strumenti automatici direttamente sulla trascrizione; di questo parleremo nel paragrafo seguente.

4.3 Interventi di *post processing*

Alcuni interventi di miglioramento, infine, possono essere attuati non tanto durante la fase di riconoscimento del parlato, ma solo sulla

9 La frequenza è indicata nei numeri che ci trovano nella colonna di sinistra di ogni tipologia di errore. Abbiamo riportato qui solo gli errori con frequenza superiore a 10.

trascrizione finale. Dai dati esposti nel paragrafo precedente abbiamo visto che la gran parte degli errori riguardano parole poco rilevanti fonologicamente ma con importante valore morfosintattico (congiunzioni, preposizioni, articoli, verbi ausiliari). Per questo motivo, si è pensato che fosse più efficace ricostruire la corretta struttura sintattica in una fase successiva, a trascrizione avvenuta.
Esistono strumenti automatici (analizzatori morfosintattici) che, utilizzando vocabolari annotati con informazioni morfosintattiche, sono in grado di analizzare un testo scritto e ricostruire la corretta struttura delle frasi ripristinando preposizioni e articoli, inserendo ausiliari laddove sia presente solo il participio ed eliminando, eventualmente, piccole parole superflue inserite erroneamente dal sistema.
Sempre in fase di post processing, è possibile intervenire su altri aspetti della trascrizione quali il trattamento dei numerali, l'inserzione di alcuni elementi interpuntivi, il controllo delle maiuscole. Naturalmente, la realizzazione di queste attività dovrà essere valutata di volta in volta, in base agli utilizzi ai quali è destinata la trascrizione. Non sempre, infatti, è più utile un testo impeccabilmente aderente al parlato.

5 Ultimi sviluppi del progetto

Attualmente le attività legate alle diverse applicazioni del sistema e soprattutto al suo miglioramento non solo non sono terminate ma sono in pieno svolgimento.
Innanzitutto, è stato ampliato il modello acustico per mezzo dell'inserzione di una maggiore quantità di audio, con un conseguente aumento del numero delle voci utilizzate per l'addestramento. In questo modo, è stato anche inserito audio con caratteristiche acustiche leggermente diverse. Sempre sul lato del modello acustico, la ricerca si sta rivolgendo all'addestramento del sistema finalizzato a 'sopportare' meglio rumori o musica di fondo; inoltre, sarà creato un modello acustico specifico per l'audio a banda telefonica.
Per quanto riguarda il modello di linguaggio, sono stati creati più modelli – oltre a quello politico, anche uno televisivo, uno religioso e uno relativo alla compilazione di telegrammi – così da poter applicare la tecnologia del riconoscimento vocale *speaker independent* a diversi ambienti. Si è visto, infatti, che la consultazione di un preciso modello

di linguaggio tarato il più possibile sull'ambiente di utilizzo del sistema contribuisce in modo significativo alla riduzione degli errori di trascrizone e al miglioramento delle *performance* del sistema stesso.
Parallelamente a tutto questo, prosegue all'interno dell'azienda l'attività di manutenzione e miglioramento dell'intero sistema, con le modalità descritte nei paragrafi precedenti.
Infine, così come era stato annunciato nel corso della nostra presentazione orale al convegno cui questi atti si riferiscono, è stato attivato un servizio di trascrizione online, sul sito www.trascrivi.it. A questo indirizzo è possibile registrarsi e inviare un file audio per averne indietro la trascrizione totalmente automatica o lavorata da un operatore specializzato dell'azienda.
Oltre a questa, molte altre sono le applicazioni possibili per questo tipo di tecnologia, dalla trascrizione di parlato finalizzata alla pubblicazione, alla sottotitolazione automatica; dalla gestione di call center, al *data mining* in archivi audio/video.
Come si è già accennato, ognuna di queste applicazioni richiede un diverso livello di accuratezza e, di conseguenza, di riduzione dell'errore. Occorre ricordare, infatti, che un sistema di riconoscimento vocale che non commettesse errori riprodurrebbe esattamente ogni elemento del parlato, mentre laddove si ha la necessità che il parlato diventi scritto, quasi mai occorre conservare nello scritto le caratteristiche del parlato. Il nostro impegno primario, dunque, perché la ricerca non rimanga fine a se stessa, non è tanto il tentativo di eliminare totalmente l'errore, quanto di far sì che il sistema produca una trascrizione il più possibile in linea con le aspettative nel singolo settore applicativo.

Appendice

	Percentuale sul totale degli errori	**Percentuale sul totale delle parole trascritte**
Sostituzioni	56,9%	9,3%
Cancellazioni	21,9%	3,6%
Inserzioni	21,1%	3,4%
Totale errori	100%	16,3%

Tabella riassuntiva dei risultati sul primo test set (circa 6 ore di parlato).

Modello di linguaggio politico (220.000.000 parole)		**Modello di linguaggio televisivo (300.000.000 parole)**		**Modello di linguaggio religioso (230.000 parole)**	
di	7766060	di	9750084	e	8385
e	4615144	e	6522003	il	6250
che	4484016	il	5682303	di	5242
il	4067928	la	5029647	la	5126
la	3702978	che	4692729	che	4330
in	2926665	a	3990565	Signore	3571
per	2917759	in	3767321	per	3404
a	2780646	per	3552384	a	2993
è	2601828	è	3417189	in	2897
del	2386536	un	3211362	del	2484
l’	2206665	l’	2742526	è	2384
un	2188813	del	2737733	non	2262
non	2166975	non	2468819	i	2103
della	1854687	i	2428655	Dio	2036
i	1750478	una	2156307	della	1835
una	1591718	della	2090485	con	1756
le	1585560	con	2041407	le	1498
si	1517716	si	1975440	l’	1452
con	1307081	le	1961920	tuo	1444
ha	1276367	ha	1773613	perché	1418
al	1139014	da	1740623	si	1414
dei	1123545	al	1495337	Cristo	1368
da	1088366	ma	1297919	nostro	1314
dell’	1031702	sono	1244348	al	1228
sono	1021275	dei	1182398	ha	1133
ma	896805	più	1127545	tua	1125
anche	892751	dell’	1080617	gli	1122
delle	857056	anche	1056944	lo	1081
alla	800585	due	1027613	un	1080
due	770840	nel	1011265	nel	1073
più	738316	alla	972258	Gesù	1029
nel	700430	gli	927085	noi	954

questo	690951	come	922187	alla	952
come	678866	delle	775893	dei	914
gli	662139	se	654479	da	868
all’	629687	all’	646758	o	867
se	600992	anni	618974	io	866
mila	532740	dal	598995	nella	864
perché	496661	lo	585548	come	850
ci	455399	mila	572394	ci	848
dal	439610	nella	540021	ma	848
questa	434338	hanno	538594	te	840
o	423704	uno	500917	loro	809
ai	416407	o	500282	tutti	803
lo	407895	tra	492230	dal	792
hanno	400517	ci	479952	egli	772
ad	399066	alle	455903	suo	764
nella	391319	ai	449368	lui	751
essere	385117	d’	447816	hai	746

Elenco delle prime 50 parole, in ordine di frequenza, riscontrate in tre diversi modelli di linguaggio.

Modello di linguaggio politico (220.000.000 parole)	
nel due mila[10]	75428
per cento </s>[11]	68338
<s> non è	51747
del due mila	51126
milioni di euro	44120

10 I numerali cardinali sono stati scomposti e convertiti in lettere, per poter essere trattati più agevolmente; verranno poi riconvertiti in cifre in fase di *post processing*.

11 I simboli <s> e </s> indicano rispettivamente inizio e fine frase.

due mila uno	38531
<s> ma non	38211
si tratta di	38066
<s> è un	37268
<s> ma il	36322
due mila due	35139
due mila tre	34865
la prima volta	33345
in grado di	32712
<s> ma la	32663
due mila quattro	32533
<s> è il	30351
<s> e poi	29464
il presidente della	29030
<s> per il	27880
due mila cinque	27503
<s> e il	27499
il due mila	27386
due mila sei	26910
<s> ci sono	26463
<s> è una	26361

per la prima	26346
<s> è la	26035
di euro </s>	25961
<s> si tratta	25876
<s> non si	25811
per cento del	25732
<s> il primo	25637
<s> il presidente	24825
una serie di	24696
<s> anche se	24604
<s> e non	24367
<s> per la	23692
<s> è stato	23675
cinque per cento	23607
la possibilità di	23605
che si è	23353
per cento di	23259
punto it </s>	22905
due mila sette	22745
<s> per questo	22298
<s> e la	22015

<s> ma è	21969
<s> il suo	21710
una sorta di	21659

Elenco dei primi 50 trigrammi, in ordine di frequenza, riscontrati nel modello di linguaggio politico.

I gesti interazionali come risorsa per la gestione dell'alternanza dei turni

Maria Eleonora Sciubba
Dipartimento di Linguistica – Università Roma Tre
sciubba@uniroma3.it

Abstract

The interactional necessity to project the next phase of an activity, and also its projection and completion, are collective actions: speakers' actions are modeled on those of the audience and orient to possible recipients' replies, expansions and insertions. This is possible thanks to the visibility of the details of multimodal actions and the participants' orientation to this visibility.
Many practices and resources are systematically employed by participants, not only to realize the transition towards the next activity phase, but also to embody their sense of completeness of the previous activity phase, and to accomplish the public availability, recognisability and interpretability of their actions.
We will analyze the way participants explicitly orient to the multimodal details of talk-in-interaction showing video-fragments and transcripts of two kinds of institutional interactions: Italian lawyer-client counseling and news interviews.
Although those videos show the complexity of speech multimodal temporality and interactional behavior, they also display participants coming, through multimodal resources, to a shared definition of a coherent and significant event. They also show that participants systematically orient through their practices to the locally available details and towards their temporal unwinding and sequential positioning.

1 Introduzione

La necessità interazionale di proiettare la fase successiva dell'attività, così come la proiezione vera e propria e il suo compimento sono un'impresa collettiva: le azioni del parlante sono modellate riflessivamente su quelle dei destinatari/ascoltatori, e si orientano verso le possibili repliche, espansioni e inserzioni di questi. Ciò è reso possibile dalla visibilità dei dettagli delle azioni multimodali e dall'orientamento dei partecipanti verso tale visibilità.
Diverse pratiche e risorse vengono sistematicamente utilizzate dai partecipanti non solo per realizzare la transizione verso la fase successiva dell'attività, ma anche per dar corpo al loro senso di

completamento della fase precedente, e per realizzare la disponibilità, riconoscibilità e interpretabilità pubblica delle loro azioni di risposta.
Per mostrare in che modo i partecipanti si orientino esplicitamente ai dettagli multimodali del parlato interazionale, useremo video-frammenti di interazioni e i trascritti di tali video. Mostreremo i dati in formati diversi: 1) nella forma di trascritti basati sulla linearità del parlato interazionale; 2) nella forma di trascritti basati sulla linea temporale a cui il parlato e le azioni interazionali si riferiscono e a cui sono sincronizzati; 3) nella forma di istantanee riprese dai video.
Sebbene i frammenti mostrino la complessità delle temporalità multidimensionali del parlato e del comportamento interazionali, essi mostrano anche che i partecipanti pervengono ad una comune definizione di un unico evento coerente e significativo attraverso le risorse multimodali, e mostrano l'organizzazione di questo evento tramite pratiche sistematiche che si orientano regolarmente verso i dettagli localmente disponibili e verso il loro svolgimento temporale ed il loro posizionamento sequenziale.
I dati analizzati provengono da corpora di interazioni istituzionali: colloqui avvocato-cliente e interviste giornalistiche.

2 Quadro di riferimento teorico

Gli studi sui *componenti costitutivi del turno* (Ford, Fox and Thompson, 1996; 2002; Ochs *et al*, 1996; Selting and Couper-Kuhlen, 2001) in Analisi Conversazionale (AC) e in Linguistica Interazionale hanno reso possibile l'identificazione delle Unità Costitutive di Turno (UCT; *Turn-Constructional Units*, *TCUs*), unità dell'organizzazione delle interazioni strutturate sfruttando le risorse linguistiche modellate sia dalle contingenze dell'organizzazione incrementale dell'interazione, sia dai vincoli e dalle possibilità offerte dalla grammatica, l'una essendo strutturata specularmente all'altra.
Le UCT sono delle unità particolari definite praxeologicamente[1] dalle azioni dei partecipanti, negoziate interattivamente, emergenti momento per momento dallo svolgimento temporale del parlato così come esso è prodotto dai partecipanti: in modo condiviso, situato e contingente. Le opportunità per partecipare e produrre interattivamente le UCT emergenti sono fornite dalla loro potenziale

[1] Si rimanda a Duranti (2003) per una sinossi sul parlato come pratica sociale.

proiettabilità (*projectability*, che chiameremo anche ***prevedibilità***), proprietà responsabile del fatto che le caratteristiche della costruzione di un turno emergente si possono prevedere. La prevedibilità (Sacks, Schegloff and Jefferson, 1974; Schegloff, 1984; Sacks, 1992) è una caratteristica centrale che fornisce al parlante successivo la possibilità, non solo di individuare il completamento del turno, ma di prevederlo prima che esso abbia luogo, e quindi che sia effettivamente prodotto. In questo modo, il destinatario può anticipare i punti di rilevanza transizionale e localizzare il punto successivo in cui iniziare a parlare (Ford, Fox and Thompson, 1996; Selting, 2000; Auer, 2005):

> [t]he sentence is a great packaging technique for a series of reasons [...]. It has a structure which can at all points be seen as to whether it is possibly complete or not possibly complete, and people are able to deal with it in such a way as to see, on its occurence that it's possibly complete. And also, from its beginning it can be looked at to see what it will take to complete it (Sacks, 1992: 40).

Inoltre,

> [s]ential constructions are capable of being analyzed in the course of their production by a party/hearer able to use such analyses to project their possible directions or completion loci (Sacks, Schegloff and Jefferson, 1974: 709).

Queste caratteristiche permettono ai partecipanti di definire i turni in maniera flessibile, dinamica, localmente gestibile, modificabile ed espandibile (Sacks, Schegloff and Jefferson, 1974: 726-7).
Le proiezioni delle azioni successive, a diversi livelli dell'organizzazione del parlato interazionale, sono una caratteristica generale del parlato molto studiata in AC. I partecipanti all'interazione mostrano il proprio orientamento ai dettagli interazionali emergenti – sia prodotti sia interpretati – monitorando costantemente la posizione sequenziale delle proprie azioni nella più ampia azione in corso. Questo aspetto costituisce la base del sistema di allocazione del turno (Sacks, Schegloff and Jefferson, 1974) permettendo ai destinatari di prevedere i punti di possibile completamento del turno – punti in cui le unità costitutive del turno probabilmente termineranno (Ford, Fox

and Thompson, 1996; 2002) – e, più generalmente, il possibile completamento delle proiezioni, delle sequenze e di più ampi frammenti di azioni.
Le proiezioni, perciò, preannunciano ciò che succederà nel momento successivo, permettendo ai partecipanti di negoziare corsi di azioni condivisi fino al momento in cui un possibile problema comunicativo sarà collaborativamente risolto. Le proiezioni si possono trovare all'interno di unità prosodiche, sintattiche (Auer, 2005), costitutive di turno (Schegloff, 1996; Selting, 2000; Ford, 2004), nelle sequenze (Drew, 1995), nei gesti (Streeck, 1995).
Inizialmente, gli studi sulle regolarità riscontrabili nella proiezione e nella costruzione incrementale delle UCT si erano incentrati principalmente sulle risorse puramente linguistiche, ma sono stati subito inclusi anche l'analisi dei gesti, degli sguardi e delle posture del corpo[2]. Nei primi studi[3] furono prese in considerazione risorse linguistiche *sintattiche*: venivano considerate primarie nella individuazione delle UCT le frasi, i sintagmi ecc. (Sacks, Schegloff and Jefferson, 1974; Sacks, 1992).
Successivamente, è stato mostrato che anche la prosodia e le risorse fonetiche giocano un ruolo fondamentale nella proiezione del turno (Sacks, 1992; Auer, 1996; Couper-Kuhlen, 1996; 2001; 2004).
Altri studi hanno, invece, cercato di prendere in considerazione non una sola dimensione per volta, rilevante per l'interazione, ma di combinare i contributi che potevano venire dalle risorse sintattiche, prosodiche, pragmatiche (Ford and Thompson, 1996; Ford, Fox and Thompson, 1996), cinesiche (Ford, Fox and Thompson, 1996; Schegloff, 1984), ottiche (Goodwin, 1981; Lerner, 2003) e

[2] Gli studi citati hanno tutti come oggetto lo studio interazionale della multimodalità: le date dei lavori mostrano come questo aspetto avesse attirato l'attenzione dei ricercatori che si occupavano di parlato interazionale già dagli anni Settanta dello scorso secolo. È da notare, infatti, che l'articolo 'Home Position', sebbene pubblicato da Sacks e Schegloff nel 2002, era stato in realtà presentato da Schegloff ad un convegno nel 1975, pochi mesi dopo la tragica morte di Harvey Sacks.
[3] Ma anche recentemente una branca della linguistica interazionale continua ad asserire il carattere logocentrico dell'interazione e la sintassi è ancora ritenuta la principale risorsa per il "confezionamento" dei turni: si pensi a Schegloff (1996), che individua nel materiale sintattico ciò che contraddistingue le posizioni iniziali, mediane e finali delle UCT, oppure ad Auer (2005), sulle potenzialità preannunciatrici della sintassi.

coinvolgenti l'intero corpo (Goodwin, 2000) per il completamento del turno.

3.1 Gesti interazionali

La necessità interazionale di proiettare la fase successiva dell'attività, così come la proiezione vera e propria e il suo completamento sono un'impresa collettiva: le azioni del parlante sono modellate riflessivamente su quelle dei destinatari/ascoltatori e si orientano verso le possibili repliche, espansioni e inserzioni di questi. Ciò è reso possibile dalla visibilità dei dettagli delle azioni multimodali e dall'orientamento dei partecipanti verso tale visibilità.
Diverse pratiche e risorse vengono sistematicamente utilizzate dai partecipanti non solo per realizzare la transizione verso la fase successiva dell'attività, ma anche per dar corpo al loro senso di completamento della fase precedente e per realizzare la disponibilità, riconoscibilità e interpretabilità pubblica delle loro azioni di risposta.
Per mostrare in che modo i partecipanti si orientano esplicitamente ai dettagli multimodali del parlato interazionale, useremo video-frammenti di interazioni e i trascritti di tali video. Mostreremo i dati in formati diversi: 1) nella forma di trascritti basati sulla linearità del parlato interazionale; 2) nella forma di trascritti basati sulla linea temporale a cui il parlato e le azioni interazionali si riferiscono e a cui sono sincronizzati; 3) nella forma di istantanee riprese dai video.
Sebbene i frammenti mostrino la complessità delle temporalità multidimensionali del parlato e del comportamento interazionali, essi mostrano anche che i partecipanti pervengono ad una comune definizione di un unico evento coerente e significativo attraverso le risorse multimodali, e mostrano l'organizzazione di questo evento tramite pratiche sistematiche che si orientano regolarmente verso i dettagli localmente disponibili e verso il loro svolgimento temporale ed il loro posizionamento sequenziale (Mondada, 2006b).

3.1.1 Chiudere il turno

Mondada (2006a) mostra che la ***proiezione della chiusura*** della fase in corso e della transizione verso la fase successiva può avvenire in diverse posizioni sequenziali. Nell'estratto ivi analizzato la studiosa

descrive tre tentativi di realizzare una transizione verso l'azione successiva, mostrando una gamma di possibilità e risorse diversamente sincronizzate: nel primo tentativo il parlante usa una pausa tra la fine di un turno e l'inizio del successivo, per chiudere una fase interazionale; nel tentativo successivo, l'interagente tende ad *anticipare* gestualmente la transizione quando la chiusura a livello di organizzazione del turno non è ancora conclusa.
Le risorse usate consistono nell'organizzazione del proprio turno finale e nel riarrangiamento dello spazio di lavoro necessario per le pratiche referenziali ed esplicative in cui i partecipanti sono impegnati. ***Muovere oggetti nello spazio*** diventa una risorsa per l'organizzazione interazionale in quanto gli oggetti sono attinenti in modo rilevante al lavoro interazionale in corso.
I destinatari organizzano varie azioni di risposta che hanno l'effetto di espandere con una coppia adiacente la sequenza che il parlante corrente intende chiudere. La varietà di risorse che essi usano per inserire la loro mossa forte è calibrata sulla temporalità e sul progetto delle azioni del parlante in corso: dall'uso di pratiche grammaticali che continuano il turno del parlante precedente (co-costruzione del turno da parte di più partecipanti; si veda anche Lerner, 1992; 1996), all'uso di deittici e di risorse multimodali per rendere nuovamente rilevante la configurazione spaziale precedente:

> so resources and practices are organized by obeying local time constraints imposed by the turn, the sequence and the bodily action, and by taking into account, within larger time unfoldings, previous solutions to the practical problem encountered, either altering or recycling them (Mondada, 2006a: 127).

Analogamente, nell'estratto che segue (1.1), l'avvocato segnala l'abbandono del piano della conversazione a favore della sua interlocutrice ricorrendo alla gestualità: nel momento in cui l'avvocato raggiunge il completamento sintattico del turno, e quindi un punto di rilevanza transizionale sintattica, la cliente approfitta dello spazio di proiezione[4] (*projection space*; Schegloff, 1984) creatosi e inizia il

[4] «The notion of a ''projection space'' is concerned with both the span in which some element of talk is ''in play'' before being produced, and with the evidence of that which a speaker's turn may make available to its recipient. For example, it appears central to the organization of tightly coordinated turn-transfer from one

proprio turno sovrapponendosi alle due ultime parole del turno precedente. L'avvocato accetta di cedere il turno alla sua interlocutrice riportando le mani sul tavolo davanti a sé (linea 12) nella posizione di partenza, o riposo (Kendon, 1996; Schegloff, 2002), ovvero sui documenti oggetto di discussione.

(1.1) **divorzio_gesto1**

8.	A;	***[6]e poi la sera andava a lav**[7]orare* [*[8]al bar*]**
9.	AvvG;	*[6].....HH ancora + dx....* *[7]..HH up..* *[8]--HH up palm up--*
10.	C;	***[8][e più] cin***
11.		***[9]que figli di cui u*no neanche**
12.	AvvG;	*[9],,,,,,,,,HH,,,,,,, --> *

Linea 9

Linea 12

Anche nell'estratto che segue (1.2), preso dalla stessa interazione, i gesti della parlante ritornano alla posizione di riposo nel momento in cui sta finendo il turno. Si può notare che la parlante inizia a ritrarre i gesti molto prima della fine del turno. In effetti, dal punto di vista dell'organizzazione sequenziale dell'interazione, questo punto segnala la fine di una sequenza e l'inizio della successiva. Anche la cornice partecipativa cambia radicalmente: l'avvocato è uscito, ed in questo momento rimangono nella stanza solo la cliente ed il fratello. Il cambio di cornice partecipativa è segnalata anche dal riposizionamento nello spazio del corpo della cliente che, chiudendo il cellulare, orienta il proprio corpo ed alza il proprio sguardo verso il

speaker to a next that aspects of some current turn are projected, and are available to analysis by a recipient/potential-next-speaker before their actual occurrence» (Schegloff, 1984: 267).

fratello. Inizia da questo momento una lunga sequenza di *code-switching*.

(1.2) **divorzio_sms**

81.	C;	**^ =c'è *il suo numero*^**
82.	CliG;	^----------pp----------^
83.	AvvG;	*si alza/si gira verso la porta*
84.		(*.3*)
85.	A;	**^*vabene sì* lei lo tenga quello= ^**
86.	AvvG;	*esce*
87.	C;	***^=sì^=sì e questo me lo so tenuto sacro***
88.	CliG;	^,,,,,tutti i gesti e posture,,,,HP // chiude il cellulare e orienta il corpo e lo sguardo verso il fratello^

Linea 88 - ^,,,,,tutti i gesti e posture,,,,HP // chiude il cellulare e orienta il corpo e lo sguardo verso il fratello^

La fine del turno può essere segnalata ancora più esplicitamente dal parlante in corso. Nell'estratto seguente (1.3), preso da un corpus di interviste giornalistiche, il parlante, alla fine del proprio turno di parola, lo cede al proprio interlocutore rilassando la postura del corpo,

chiudendo gli occhi e sorridendo. Tutte queste azioni avvengono simultaneamente.

(1.3) **Epifani_mezzora**

1.	E;	**mentre farlo in forma- (.) volontaria di incentivazione di creare**
2.		**d=ll=condizioni di restare di più a lavoro (.) non abbiamo mai avuto**
3.		**obiezioni.**
4.	EpifG;	§-----§[2]----------§ ((*postura si rilassa//occhi chiusi//sorride*))
5.		§(.)§
6.	A;	**§[2]h e§cco però epifani**

Linea 4

Concludendo, con lo scorrere del tempo e lo svolgersi dell'interazione i partecipanti mostrano la loro condotta verbale e corporea in modo da proiettare altre azioni; questa esibizione è orientata verso il destinatario e la cornice di partecipazione viene reciprocamente adattata, nel corso dell'azione, alla propria ricezione incarnata (*embodied*) di tale esibizione. Perciò, perché l'interazione possa andare avanti, devono essere esplicitamente disponibili e visibili in anticipo le risorse interazionali. In questo senso anche l'analisi *in tempo reale* lega riflessivamente la produzione e l'interpretazione dell'azione, quest'ultima essendo contenuta nella prima (Goodwin, 1981). In altre parole, sono disponibili non solo i piani sistematici dei turni di parola, ma anche altre traiettorie multimodali preannunciate e la "configurazione contestuale" che gli artefatti, le azioni e l'organizzazione spaziale dei partecipanti e degli oggetti formano in modo contingente e rilevante. Queste configurazioni non sono distribuzioni statiche di oggetti nello spazio, ma cambiano costantemente nel tempo durante cui l'azione si svolge, termina o si trasforma: stiamo parlando di ciò che Goodwin (2000: 1491) chiama la *geography of cognition*, la geografia cognitiva situata sia nell'intersoggettività dell'organizzazione sequenziale di parlato ed

azione sia nella disposizione dinamica del contesto fisico/materiale e sociale.

3.1.2 Prendere il turno

Mondada (2006b) mostra che le caratteristiche dell'ambiente materiale e spaziale circostante sono costantemente "mobilitate" per ottenere e realizzare l'autoselezione ed anche per altre pratiche di presa del turno.
Analizziamo le pratiche di presa del turno osservabili in un corpus particolare, affrontato come se fosse un *setting perspicuous* (Garfinkel and Wieder, 1992), contesto evidente, ovvero un setting unico e specifico che ci permette di evidenziare caratteristiche sistematiche ed ordinate. Questo setting offre risorse multimodali specifiche, connesse alla disposizione spaziale dei partecipanti, agli oggetti che manipolano e alle loro particolari attività.
Le attività considerate sono caratterizzate da un'ecologia particolare, in cui uno spazio di lavoro limitato, pieno di oggetti mobilizzati dal parlato e dai gesti, è al centro dell'attenzione collettiva. Contrariamente ad altri contesti di lavoro in cui attività multiple fanno costantemente disperdere l'attenzione in spazi frammentati, e in cui l'attenzione reciproca deve essere costantemente riguadagnata tramite un intenso lavoro interazionale (Goodwin, 2003; Heath *et al*, 2002), l'attività qui studiata è caratterizzata da un focus sostenuto di attenzione collettiva sui documenti posti sul tavolo. Questa particolare ecologia offre risorse specifiche per l'organizzazione dell'interazione, sia per quanto riguarda la deissi sia, soprattutto, per l'organizzazione sistematica dell'interazione.
Nei dati video analizzati in Mondada (2006b) è molto ricorrente, da parte dei partecipanti all'interazione, il gesto di indicare (*pointing*) le piantine degli edifici e gli altri documenti presenti sul tavolo di lavoro. I gesti di indicare sono stati analizzati in letteratura principalmente in relazione alla deissi (Hanks, 1992; Kendon, 2004): sono una risorsa efficace e pervasiva per identificare un oggetto e renderlo riconoscibile, anche se non è visibile ai partecipanti (Haviland, 1996), all'interno di pratiche referenziali che possono o meno cooccorrere con il parlato. Studi sul gesto di indicare nelle interazioni (Hindmarsh and Heath, 2000; Goodwin, 2003) hanno mostrato che l'articolazione

dei deittici e dell'indicare non è così semplice come può sembrare: i gesti di indicare sono perfettamente sincronizzati con l'organizzazione momento per momento del parlato interazionale, con il parlato orientato al destinatario ed il comportamento del corpo, con appropriate disposizioni dei corpi e degli oggetti nello spazio, con il progressivo esame minuzioso del contesto e di ciò che avviene, con i bisogni contingenti dell'attività in corso e con la coordinazione di tutto ciò opportunamente organizzata.
Inoltre, l'indicare non è solo ed esclusivamente associato alla deissi: nel corpus analizzato in Mondada (2006b) questi gesti sembrano connessi, in prima analisi, alle pratiche descrittive che caratterizzano questa attività lavorativa, ma anche alla ridefinizione dell'attenzione comune verso dei referenti condivisi.
I partecipanti indicano le piantine non solo per riferirsi ad esse ma anche per rendere visibile il loro coinvolgimento e partecipazione all'interazione. La posizione temporale e strutturale dell'indicare è una risorsa chiave per l'organizzazione della presa del turno. Le pratiche per la presa del turno e per indicare sono qui strettamente connesse, in un modo che sfrutta le particolarità dell'azione situata: ovvero, essa avviene intorno ad un tavolo coperto di oggetti, l'attenzione dei partecipanti è focalizzata più su tali oggetti che sui partecipanti stessi e le attività locali sono organizzate intorno a oggetti concreti.
L'analisi concerne due aspetti dell'indicare: il contesto sequenziale in cui il gesto inizia, che caratterizza il punto in cui il destinatario/ possibile parlante successivo sta emergendo e si sta caratterizzando come parlante; e il contesto sequenziale in cui il gesto termina, che non coincide necessariamente con la fine del turno del parlante. Perciò l'indicare è utilizzato dai partecipanti per autoselezionarsi come parlante successivo e per preannunciare l'imminente completamento del turno del parlante attuale (ivi).
Questo fine descrittivo pone dei vincoli stretti sui dati utilizzati (Mondada, 2006a): da un lato richiede che i dati siano videoregistrati in modo da rendere accessibili al ricercatore i dettagli rilevanti; dall'altro, comporta trascrizioni che rappresentino la fine sincronizzazione dei dettagli multimodali nel momento in cui si dipanano durante l'interazione, rendendoli disponibili all'analisi.
La studiosa (ivi) giunge alle seguenti conclusioni:

(1) Per la descrizione del gesto, la posizione sequenziale dell'inizio della sua fase preparatoria è essenziale e permette l'identificazione di un contesto sequenziale rilevante per il fenomeno.
(2) Il gesto di indicare non è limitato alla mano: coinvolge l'intero corpo (ad esempio, la parte superiore del corpo che si piega sul tavolo, il braccio che attraversa il tavolo, v. es. 1.6). Inoltre, l'incassamento dell'indicare in altre attività e nell'ecologia dell'azione in corso modella la forma e le risorse del gesto: indicare con una matita o indicare con un dito permettono tipi diversi di movimenti, di precisione e di attività parallele che possono avvenire simultaneamente, a seconda della parte del corpo utilizzata.
(3) Il gesto dell'interagente non è un mero gesto individuale isolato, ma un gesto finemente coordinato, sincronizzato con il formato del turno in corso e adattato alle azioni degli altri.

Come già ricordato, i gesti sono potenzialmente anticipatori: non solo i gesti iconici precedono gli elementi del parlato a cui sono associati, ma più in generale, i gesti sono orientati verso l'organizzazione dell'interazione e preannunciano posizioni sequenziali rilevanti, come ad esempio i punti di rilevanza transazionale. In Mondada (2006a), l'indicare non precede solo correlati deittici, ma è sfruttato più genericamente per preannunciare l'autoselezione (Streeck, 1995, parla di preannunciatore di azioni, *action projector*); compare non solo nei PRT ma tende ad iniziare anche prima, aprendo uno spazio di proiezione (Schegloff, 1984).
Riassumendo, l'indicare si verifica intorno ai punti di rilevanza transazionale, sebbene possa verificarsi molto prima; si presenta come un *metodo* per preannunciare l'autoselezione, essendo parte di un movimento emergente che serve per stabilire l'imminente formato di parlante; rende visibile a tutti i partecipanti un pre-inizio, invadendo il campo visivo ma non ancora quello vocale del turno.
Questa fase di pre-inizio anticipa il possibile completamento dell'unità che è prodotta dal parlante attuale: a) mostra il monitoraggio della struttura in corso da parte del possibile parlante successivo e, più precisamente, la sua analisi finemente armonizzata per tutti i fini interazionali possibili; b) preannuncia l'azione successiva, ovvero, il turno che sarà preso dal destinatario/prossimo parlante che sta indicando.
Perciò, l'indicare ha più a che vedere con la sincronizzazione della costruzione delle UCT *in fieri* che con le pratiche referenziali *tout court*, dal momento che l'indicare inizia molto prima che un deittico

venga pronunciato, o addirittura in assenza di deittici, ovvero, si verifica molto prima la rilevanza di indicare deitticamente qualcosa. Ciononostante, se osserviamo le posizioni in cui in questi enunciati vengono prodotti i deittici, possiamo osservare che il riferimento deittico in posizione ad inizio di turno può rappresentare una risorsa per aumentare la legittimità e la responsabilità dell'indicare come metodo per riformulare la cornice partecipativa e stabilire il ruolo di parlante (Mondada, 2006b).

(1.4) **Epifani_mezzora**

1.	A;	§**siamo**§ §[2]**a pericolo:**§
2.	EpifG;	§mani incrociate§§[2]apre le mani H§
3.	A;	§**e [ritorniamo sul] tema in #questa s#ituazione**§
4.	EpifG;	§----------H----------§
5.	E;	[h_.8]
6.	AnnG;	#pp -> E#
7.	A;	§**che lei descrive di una nuova ondata h ^↑addirittura terroristiche↑^ c'è**§
8.	EpifG;	*ppp----------> *
9.	EpifS;	^-----alza sguardo----------^ >> (mantiene fine t)
10.	A;	§**chi dice che si parla di nuovo terrorismo**§**.**
11.	EpifG;	§ppp,,,,,,,,,,,,,,, -> §

Linea 2 ((apre le mani))

Linea 8 ((indica))

Nell'estratto precedente, è evidente come il gesto di indicare sia rilevante per la presa del turno: Epifani muove le mani per segnalare la sua intenzione di prendere la parola (l. 2; 8), ed il mantenimento del gesto di *puntamento* (l. 6) segnala per tutto il tempo della sua durata che questo proposito perdura. L'intenzione di Epifani (sempre frustrata dall'intervistatrice) di prendere il turno è anche mostrata dalla lunga inalazione (5 decimi di secondo) alla linea 5.

3.1.3 Mantenere il turno

I partecipanti all'interazione che ricoprono il ruolo di destinatari, e che preannunciano la propria autoselezione attraverso le risorse multimodali, mostrano anche un monitoraggio sintattico dei turni dei parlanti in corso, con il fine pratico di cercare punti di possibile rilevanza transazionale. Ad esempio, la sospensione dei gesti mostra il riconoscimento non solo degli spazi di proiezione ma anche della loro ampiezza: quando il punto di proiezione è rimandato, il gesto si rilassa e viene sospeso; quando si riavvicina un punto di proiezione, il gesto viene ripreso. Questo dimostra che i partecipanti si orientano verso la rilevanza e la legittimità, in alcune posizioni sequenziali, della loro rivendicazione del turno di parola (come emerge dall'estratto seguente 1.5), ma si può notare anche che, in altri momenti, queste rivendicazioni non possono essere sostenute e devono essere ritirate o abbandonate del tutto (Mondada, 2006b), come si è visto nell'estratto precedente (1.4).

(1.5) **divorzio_sms**

68.	C;	***questo è il diciotto di ottobre alle sei e cinquanta mi stanno facendo una perquisizione non posso portare *flaviana* chiedile scusa da parte mia***
69.	Avv G;	*----------piegato sul cellulare // legge----------*
70.		**(.6)**
71.	A;	**°e risulta che- da dove proveniva la chia^mata- il messa^^2 ggio?°**
72.	CliG;	^.....pp.....^ ^2----pp---^ >>>
73.		**(.)**
74.	C;	**^2 ebbè dal suo numero**
75.		**(.3)**
76.	A;	**ah vabbè=**
77.	C;	**=sì=sì come no^ ^3dal suo- eccolo guar^di***
78.	CliG;	^3,,,,,ppp,,,,^
79.		**(^^1^^)**
80.	CliG;	^....pppp....^ ^2----pppp----^
81.	A;	**^2ah vabbè=^**
82.	C;	**^ =c'è *il suo numero*^**
83.	CliG;	^----------pp----------^

Linea 72; 79; 82

In questo caso, grazie al mantenimento del proprio gesto di indicare, il parlante tratteggia la durata della propria azione, coprendo più turni consecutivi. In particolare, la cliente sta guidando l'avvocato nella lettura di un messaggio testuale sul proprio cellulare che lei, ovviamente, conosce già.

Nel brano seguente si assiste ad un'azione simile, in cui la cliente per tutta la durata della propria azione mantiene la posizione sporta sul tavolo, indicando (gesto 2) o giocherellando con i bigliettini da visita (gesto 3) nell'esatta posizione in cui avviene lo scambio dei bigliettini fra cliente ed avvocato. Questo mantenimento del gesto è anche un modo di rimanere nello spazio visivo periferico dell'avvocato che è chino sui documenti su cui sta prendendo appunti.

Alla linea 14 inizia l'azione di porgere i bigliettini, che vengono presi dall'avvocato alla linea 17. Fintanto che l'avvocato ha in mano i bigliettini la cliente rimane sporta occupando lo spazio centrale del tavolo, incluso nella visione periferica dell'avvocato. Alla linea 24, la cliente riprende i biglietti e continua a mantenere la sua posizione sporta fino alla linea 27, in cui inizia a ritornare alla posizione iniziale (*home position*, Schegloff, 2002).

Come si può notare, la cliente ritorna in posizione iniziale nel momento in cui ha finito di riferire tutte le informazioni che ritiene che l'avvocato debba avere, ovvero quando reputa che la sequenza debba terminare.

(1.6) **20080122divorzioGesto2bigliettini**

14.	F;	insomma:::=mm ***** [[2]>stavo in zona molti anni fa<]
15.	A;	***** [[2]tanto quello poi quello] lo possiamo- lo possiamo-
16.	F;	[(°lei crede°)]
17.	C;	[vede avvocato] **§** questi erano <i suoi due> al di là del bar (.3)
18.		questi [erano] i due esercizi
19.	A;	[sì]
20.	C;	che [lui aveva aperto:]
21.	A;	[e infatti che lui] dice::-
22.		che [lui di]ce che aveva abbandonato
23.	C;	[ecco]
24.	A;	[per]ché lei [[2] non l'aiutava] **$** no?
25.	C;	[eh:] [[2] eh vabbè]
26.		(.4)
27.		**£** eh ma il <u>bar:</u> (.) e:: non era <u>lui</u> proprietario eh ma era in condivisione**"**con la moglie
28.		(.7)
29.	A;	°[ah con l'ex] [[2] moglie]°
30.	C;	[che lui parl-] [[2] ↑eh sì↑]
31.		(1.3)
32.	A;	°allora (.) aspetti un attimo xxxxxxxxxxxxx°
33.		((inizia a scrivere))

***** Gesto 1 (porge i biglietti)

§ Gesto 2 (indica i biglietti)

$ Gesto 3 (riprende i biglietti)

£ Gesto 4 (mantiene la posizione)

“ Gesto 5 (Home position)

Questi dati suggeriscono che i partecipanti, ed in particolare i parlanti, organizzano il raggio d’azione dei propri turni, coprendo non solo la durata del turno prodotto, ma la sequenza – o più sequenze successive – che essi iniziano. In questo modo, i gesti dei parlanti rivelano che i diritti e gli obblighi che caratterizzano il formato di produzione di parlante non finiscono con la fine dei turni, ma continuano insieme alle aspettative relative alla sequenza iniziata con quel turno. Perciò, il parlante rimane “responsabile” della sequenza che ha iniziato e continua a controllarla, manifestando se è completa o se è stata recepita adeguatamente (Mondada, 2006b).
Il fatto che questa risorsa interazionale sia comunemente impiegata, si evince anche dal seguente estratto preso dal corpus di interazioni istituzionali di interviste giornalistiche.

(1.7) **Epifani_mezzora**

1.	E;	**e quindi non credo che lui si riferisca [a questo]**
2.	A;	***[però] è un ragionamento diciamo diverso sul sud è un ragionamento di: controllo***
3.	EpifG;	*((*mano dx aperta sul tavolo con indice verso A e l’alto*))----------*
4.	A;	***dei fondi è un ragionamento sulla qualità***
5.	EpifG;	**porta l’indice della mano sx alla punta del naso**

Linea 3

Linea 5

Epifani, in questo caso, utilizza il mantenimento di un gesto come risorsa conversazionale per segnalare all'interlocutore che gli ha rubato il turno che non vi ha rinunciato. Quando si accorge che la giornalista, e quindi regista dell'interazione, si è imbarcata in un turno lungo e, soprattutto, sta utilizzando i suoi diritti di regia sull'interazione stessa, reindirizzando la risposta dell'intervistato sulla sua interpretazione dei fatti, Epifani ritira il gesto interazionale, con cui voleva indicare di volersi riappropriare del turno, e con l'indice dell'altra mano si tocca la punta del naso. Anche quest'ultimo può essere considerato un gesto interazionale: segnala all'interlocutore che il parlante precedente non vanta più pretese sul turno appena usurpato.

4 Conclusioni

Gli studi sul parlato interazionale si basano su una prospettiva di studio definita *praxeologica* che colloca la cognizione non nella testa di un soggetto isolato, ma nella produzione ordinata e riconoscibile di azioni nel momento in cui esse sono progettate, affrontate e, se necessario, riparate dai partecipanti alle interazioni (Schegloff, 1991; 1992): le azioni e la responsabilità per tali azioni sono costruite interattivamente dagli interagenti; per produrre ed interpretare azioni riconoscibili, i partecipanti si orientano verso i dettagli che costituiscono l'ordine locale di parlato ed azione e mostrano reciprocamente il loro orientamento attraverso il loro comportamento; questi dettagli esplicitamente visibili riguardano varie risorse e pratiche multimodali che sfruttano sistematicamente le caratteristiche del linguaggio, i gesti, gli sguardi, i movimenti del corpo e la manipolazione di oggetti.

Il fatto che tali gesti "interazionali" sembrano essere prodotti, negli esempi riportati, soprattutto da coloro che hanno minor potere interazionale, e quindi non possono accedere alle risorse di gestione dei turni riservate al regista dell'interazione (Orletti, 2000), hanno portato ad ipotizzare che la funzione di tali gesti sia quella di evitare possibili insubordinazioni ai poteri del regista che possono essere sanzionate. In questo modo la stessa nozione di prevedibilità/proiettabilità degli elementi costitutivi dei turni aiuterebbe il regista ad adempiere al suo compito di "conduttore" dell'interazione, senza dover decidere se il turno di parola autoselezionato costituisca insubordinazione o meno. Questa ipotesi,

per l'esiguità dei casi esaminati, non deve essere presa come generalizzazione di comportamenti, ma come indizio per future indagini nel campo delle interazioni istituzionali e fra pari.

Appendici

A1 – Sistema notazionale

SIMBOLI	FENOMENI	COMMENTI
parola parola	parole	spazio prima e dopo i confini di parola
	alternanza dei parlanti	inizia una nuova riga quando avviene un cambio di parlante
	sequenza dei turni	la sequenza dell'alternanza dei parlanti marca approssimativamente la sequenza dei turni
1	numero di riga	ciascuna riga è contraddistinta da un numero progressivo
PM;	attribuzione del turno	il punto e virgola segue il nome o pseudonimo del parlante
...	pausa, non calcolata	pausa > 0.2 secondi
	vocalizzazioni	uh, um, mm, unh-unh, eh ecc. (ortografia onomatopeica)
@ haha	risata	una per ciascuno scoppio o particella di risata
[ ]	sovrapposizione (prima serie)	le parentesi sinistre sono allineate verticalmente
[2]	sovrapposizione (seconda serie)	le parentesi sono allineate verticalmente, co-indicizzate con un pedice progressivo
((commento))	commenti	
paro–	parola troncata	
xxxx	segmento incomprensibile	segmento o parola incomprensibile
(.h)	inspirazione	
(h)	espirazione, sospiro	
(1.2)	pausa, calcolata	durata della pausa in secondi e decimi di secondo
:	allungamento	i due punti rappresentano l'allungamento di una vocale o consonante
..	micropausa	< 150 millisecondi
=	allacciamento	due parole consecutive sono pronunciate senza pausa fra di loro
(X)_(n.n)	durata (opzionale)	tempo in secondi
(TSK)	click	click alveolare
%	colpo di glottide, creak (voce stridula)	vocalizzazione separata = "parola" separata

MAIUSCOLO	volume	le parole più forti, o urlate, rispetto al resto sono scritte in maiuscolo
testo	enfasi	può riguardare tutta la parola o parte di essa
↓parola↓	piano, attenuato	
> <	parlato veloce	
< >	parlato rallentato	
# #	parlato scandito	

A 2 Gestualità

FENOMENI	SIMBOLI
* *	Indicano l'inizio e la fine di un gesto dell'avvocato, sovrapposto al parlato
*2 *	Indicano l'inizio e la fine di un gesto dell'avvocato, sovrapposto al parlato. Il numero indica la sequenzialità della sovrapposizione (seconda, terza, ecc., serie)
^ ^	Indicano l'inizio e la fine di un gesto della cliente, sovrapposto al parlato
^2 ^	Indicano l'inizio e la fine di un gesto della cliente, sovrapposto al parlato. Il numero indica la sequenzialità della sovrapposizione (seconda, terza, ecc, serie)
…	preparazione o inizio del gesto
---	apice o colpo (stroke), o mantenimento del gesto
,,,,	fine/ritiro del gesto, ritorno alla posizione di riposo
ppp,,,,,,,,,,,,,,, ->	Continuazione del gesto fino alla fine del turno
ppp,,,,,,,,,,,,,,, ->>	Continuazione del gesto fino alla fine del turno per più righe, o per più turni
HH	entrambe le mani
H	una mano
pp	gesto di indicare
HP	Home Position Posizione di inizio a cui ritorna il gesto

Riferimenti bibliografici

Couper-Kuhlen, E. (1996) The prosody of repetition: On quoting and mimicry, in E. Couper-Kuhlen and M. Selting (eds) *Prosody in Conversation: Interactional studies*, Cambridge: Cambridge University Press, pp. 366-405.

– (2001) Intonation and discourse: Current views from within, in D. Schiffrin, D. Tannen and H. Hamilton (eds) *Handbook of Discourse Analysis*, Oxford: Blackwell, pp. 13-34.

– (2004) Prosody and sequence organization: The case of new beginnings, in E. Couper-Kuhlen and C.E. Ford (eds) *Sound Patterns in Interaction*, Amsterdam: Benjamins, pp. 335-76.

Duranti, A. (2003) Il parlare come pratica sociale, in G. Mantovani (ed.) *Manuale di psicologia sociale*, Firenze: Giunti, pp.44-61.

Ford, C.E., B.A. Fox and S.A. Thompson (1996) 'Practices in the Construction of Turns: The 'TCU' Revisited'. *Pragmatics* 6, 3, pp. 427-54.

Ford, C.E. and S.A. Thompson (1996) Interactional units in conversation: syntactic, intonational, and pragmatic resources for the management of turns, in E.A. Schegloff, E. Ochs and S.A. Thompson (eds) pp. 134-84.

Goodwin. C. (1981) *Conversational Organization: Interaction Between Speaker and Hearers*. New York: Academic Press.

– (2000) 'Action and Embodiment Within Situated Human Interaction'. *Journal of Pragmatics* 32, pp. 1489-522.

– (2003) Pointing as Situated Practice, in S. Kita (ed.) *Pointing: Where Language, Culture and Cognition Meet*, Mahwah, NJ: Lawrence Erlbaum, pp. 217-41.

Hanks, W.F. (1992) The Indexical Ground of Deictic Reference, in A. Duranti and C. Goodwin (eds) *Rethinking Context, Language as an Interactive Phenomenon*, Cambridge: Cambridge University Press, pp. 43-77.

Haviland, J.B. (1996) Projections, transpositions, and relativity, in J.J. Gumperz and S.C. Levinson (eds) *Rethinking Linguistic Relativity*, Cambridge: Cambridge University Press, pp. 271-323. Available on-line

http://www.anthro.ucsd.edu/~jhaviland/Publications/HavilandTranspositionsPublished.pdf (accessed: 27 September 2009).
Hindmarsh, J. and C. Heath (2000) 'Embodied references: A study of deixis in workplace interaction'. *Journal of Pragmatics* 32, pp. 1855-78.
Kendon, A. (1996) 'An agenda for gesture studies'. *The Semiotic Review of Books* 7, 3, pp. 7-12.
– (2004) *Gesture: Visible Action as Utterance*. Cambridge: Cambridge University Press.
Lerner, G.H. (1992) 'Assisted storytelling: Deploying shared knowledge as a practical matter'. *Qualitative Sociology* 15, 3, pp. 247-71.
– (1996) On the "semi-permeable" character of grammatical units in conversation: Conditional entry into the turn space of another speaker, in E. Ochs, E.A. Schegloff and S.A. Thompson (eds) *Interaction and Grammar*, Cambridge: Cambridge University Press, pp. 238-76.
– (2003) 'Selecting next speaker: The context-sensitive operation of a context-free organization'. *Language in Society* 32, 2, pp. 177-201.
Mondada, L. (2006a) La pertinenza del dettaglio. Registrazione e trascrizione di dati video per la linguistica internazionale, in Y. Bürki and E. De Stefani (eds) *Trascrivere la lingua. Dalla filologia all'analisi conversazionale*, Bern: Lang, pp. 313-44.
– (2006b) Video Recording as the Reflexive Preservation and Configuration of Phenomenal Features for Analysis, in H. Knoblauch, J. Raab, H.G. Soeffner and B. Schnettler (eds) *Video Analysis*, Bern: Lang, pp. 51-67.
Orletti, F. (2000) *La conversazione diseguale*. Roma: Carocci.
Sacks, H. and E.A. Schegloff (2002) 'Home position'. *Gesture*, 2 (2): 133-146. Available on-line from http://www.sscnet.ucla.edu/soc/faculty/schegloff/pubs/ (accessed: 27 September 2009).
Sacks, H., E.A. Schegloff and G. Jefferson (1974) 'A Simplest Systematics for the Organization of Turn-Taking for Conversation'. *Language* 50, 4: pp. 696-735. Available on-line from http://www.sscnet.ucla.edu/soc/faculty/schegloff/pubs/ (accessed: 27 September 2009).

Schegloff, E.A. (1984) On some gestures' relation to talk, in M.J. Atkinson and J. Heritage (eds) *Structures of Social Action. Studies in Conversation Analysis*, Cambridge: Cambridge University Press, pp. 266-296. Available on-line from http://www.sscnet.ucla.edu/soc/faculty/schegloff/pubs/ (accessed: 27 September 2009).

– (1991) Conversation Analysis and Socially Shared Cognition, in L. Resnick, J. Levine and S. Teasley (eds) *Perspectives on Socially Shared Cognition*, Washington, D.C.: American Psychological Association, pp.150-171. Available on-line from http://www.sscnet.ucla.edu/soc/faculty/schegloff/pubs/ (accessed: 27 September 2009).

– (1992) 'Repair After Next Turn: The Last Structurally Provided Defense of Intersubjectivity in Conversation'. *American Journal of Sociology* 97, 5, pp. 1295-345. Available on-line from http://www.sscnet.ucla.edu/soc/faculty/schegloff/pubs/ (accessed: 27 September 2009).

Streeck J. (1995) On Projection, in E. Goody (ed.) *Interaction and Social Intelligence*, Cambridge: Cambridge University Press, pp.84-110.

La costruzione di una rappresentazione condivisa attraverso l'impiego del focus group

Patrizia Velotti, Cecilia Serena Pace, Giulio Cesare Zavattini
Department of Dynamic and Clinical Psychology
patrizia.velotti@uniroma1.it

Abstract

The needs identified in cross-cultural research, such as to give greater attention to the processes of assimilation of cultural values, beliefs and norms (Leventhal, Musumeci and Contrada, 2007) can be pursued through studies that investigate the communication processes between individuals from different cultures. The focus group is one of the qualitative techniques used in the field of applied social research. In this type of group, each participant becomes aware of the "views" of others on this issue and discusses them in the debate, triggering a process of negotiation that makes the data collected the product of interpersonal exchanges, rather than just expression of individual perspectives (Sondika et al*, 2000). The aim of this work is to analyze the transcripts of focus group, attended by individuals from different country, which have as their object of deepening the concepts of health. The groups are made up of a sample of 30 nurses: 10 from Eastern Europe, 10 from South America and 10 from Italy. Were used the following tools: 1) the interview for the collection of personal data; 2) n. 5 focus group, three intra-ethnic and two inter-ethnic relations. Focus groups were videotaped and fully transcribed. Text analysis has allowed a careful reflection on the communication that has developed and the changes that have occurred during the meetings. At first sight the fundamental differences seem to relate to the perception that each group had of issues that affect health.*

> *«Cosa significano per noi la salute e la malattia? Cosa sono per un individuo appartenente alla nostra società? Dall'esperienza personale di ciascuno, ma anche attraverso l'informazione che circola nella società, modellata dai suoi stessi valori, le immagini si strutturano: essere malati, essere in buona salute, sono delle immagini nate primariamente da un'esperienza individuale che non può essere condivisa» (Herzlich, 1973, p. 11).*

1 Introduzione

I cambiamenti demografici degli ultimi decenni hanno aumentato le conoscenze di ciascuno circa gli altri popoli, ambienti e costumi, ma hanno anche evidenziato la necessità di esaminare, attraverso studi cross-culturali, l'influenza che questi aspetti hanno sui soggetti che vivono in contesti diversi da quelli di appartenenza (Fortier and Bishop, 2003). In Europa, in particolare, la presenza di migranti impiegati in numerosi servizi sanitari, pubblici e privati, ha reso questa variabile un dato di ricerca ancor più ricco d'interesse (Aranda and Knight, 1997). Gli studiosi hanno, infatti, sottolineato l'esigenza di esaminare gli effetti della crescente diversità etnica sulla possibilità di condivisione e scambio che questi nuovi contesti di cura "condivisi" creano (Valle, 1998; Janevic and Connell, 2001; Velotti and Zavattini, 2008).
Il processo di rielaborazione personale dell'esperienza migratoria, inoltre, è stato il focus di un interessante filone di ricerche che, in particolare, ha tentato di mettere in evidenza gli aspetti in grado di facilitare un percorso positivo di rielaborazione oppure ostacolarlo, aprendo la strada a situazioni di disagio psicologico ed esistenziale (Kurman and Ronen-Eilon, 2004; Bakker, Van Oudenhoven and Van Der Zee, 2004; Portes, Zady and Dunham, 1998). Nonostante le ricerche non sempre siano state condotte con campioni di migranti veri e propri, quanto piuttosto con campioni di soggiornanti stranieri che trascorrono periodi temporanei all'estero per motivi di lavoro o di studio, nel complesso i risultati possono definire, seppure orientativamente, i principali fattori in grado di facilitare o al contrario ostacolare un'elaborazione positiva del processo migratorio. Questi risultati mettono in luce come la possibilità di accedere ed impadronirsi delle competenze idonee ad interagire nel nuovo contesto culturale sia un elemento fondamentale, capace di facilitare il percorso di adattamento sociale (La Framboise, Coleman and Gerton, 1993; Ward and Styles, 2003): si tratta dell'acquisizione di competenze di ordine linguistico, relazionale e sociale, che supportano e rendono possibile l'interazione e il dialogo nel paese di accoglienza. Kurman e Ronen-Eilon (2004) hanno a questo proposito specificato la rilevanza del grado di conoscenza che il migrante giunge ad avere rispetto alle convenzioni sociali condivise dagli abitanti del nuovo contesto e del grado di conoscenza dei valori che regolano le condotte della società ospitante.

Lo studio dei processi di acculturazione rappresenta il focus di un'ampia parte della letteratura inerente al tema della migrazione. I *processi di acculturazione* riguardano l'incontro dei migranti con i valori e la cultura della società di accoglienza; essi infatti, come sostiene Berry (2008), hanno inizio dal momento in cui due gruppi di individui con differenti background entrano in contatto, e conducono a importanti cambiamenti culturali e psicologici per ambedue le parti. Prendendo in considerazione le strategie e le scelte che in linea generale guidano i migranti nell'interazione con il nuovo paese, è stato possibile osservare che

> le strade che coloro che migrano hanno a disposizione per relazionarsi con il nuovo contesto sociale sono molteplici e rendono conto di un ampio spettro di possibilità di regolazione dei propri sistemi di riferimento valoriali (Berry, 1997: 17).

Berry, infatti, oltre a ribadire la *mutualità* che governa i processi di acculturazione, sottolinea il fatto che essi non conducono solo a processi di *assimilazione*, ovvero di omogeneizzazione della cultura dell'individuo o del gruppo minoritario con la cultura della società "dominante" di accoglienza (Berry, 2008). Esistono, infatti, anche processi di *separazione*, ovvero di rigido rifiuto di rinunciare alle proprie consuetudini o regole culturali, per quanto banali o importanti, assieme all'evitare le interazioni e i contatti con persone/gruppi appartenenti alla società ospitante; di *marginalizzazione*, quando l'individuo o il gruppo di migranti ripone un interesse minimo sia verso il mantenimento delle proprie peculiarità culturali, sia verso quelle delle società ospitante; di *integrazione*, situazioni in cui l'individuo migrante o il gruppo, nonostante l'interesse per il mantenimento del proprio patrimonio culturale, mostrano apertura nei confronti delle persone e dei valori diffusi nel paese di accoglienza. Riuscire a conservare le caratteristiche distintive del proprio bagaglio culturale, senza precludere la possibilità di scambio e interazione con la cultura ospitante e il suo sistema di riferimento valoriale, rappresenta secondo l'autore l'opzione che permette in misura maggiore di stabilire connessioni e legami tra aspetti ed elementi delle diverse culture per ricomporle in una cornice che le contenga entrambe (Berry, 2008).

2 Il “focus group” come strumento di indagine

Il focus group può essere definito come una tecnica qualitativa di rilevazione dei dati fondata sulle informazioni che emergono da una “discussione di gruppo” su un tema o su specifici aspetti di un argomento che il ricercatore intende indagare in maniera approfondita (Zammuner, 2003). La caratteristica principale, quindi, di questo strumento (Palumbo, 2001), che poi risulta essere anche il pregio della metodologia, sta nell’*interazione* che si crea fra i partecipanti, interazione che produce idee in misura assai maggiore rispetto ad esempio all’intervista singola, sia a livello di quantità che a livello di qualità di approfondimento (Stagi, 2000).

> Le tecniche di gruppo incorporano nel loro interno i processi mentali e comunicativi che gli attori sviluppano a partire da una conoscenza condivisa dei dati del contesto (Palumbo, 2001: 168).

Anche secondo Corrao (2000), il grande vantaggio dell’interazione è quello di riprodurre in modo più realistico il processo che presiede alla formazione delle opinioni e come ciò avvenga in misura maggiore quando si ha modo di interagire *conversando* in maniera personale e meno strutturata, proprio come nel focus group.
Le informazioni fornite dai partecipanti durante la discussione costituiscono i “dati” del focus group: questa peculiarità lo rende dunque un metodo particolarmente utile per l’esplorazione approfondita delle opinioni, degli atteggiamenti e dei comportamenti della collettività e per l’approfondimento delle motivazioni sottostanti al pensiero ed al comportamento umano (Zammuner, 2003). L’utilità di questo strumento è data dalla possibilità di cogliere l’*interazione*: diversamente dall’intervista, infatti, i partecipanti non hanno il compito di rispondere individualmente, ma quello di discutere *tra* loro di uno specifico argomento, e ciò permette di ottenere risultati difficilmente raggiungibili con altre metodologie (Zammuner, 2003). Di fatto, come spiega Flick

> in linea generale il focus group può essere utilizzato come una simulazione di discorsi e conversazioni quotidiani e si può intendere come un metodo “quasi-naturalistico” utile per lo studio delle rappresentazioni e della conoscenza sociale (Flick, 1998: 59)

ed a nostro avviso del *linguaggio*.

Di recente, infine, la ricerca in *ambito sanitario*, si è avvalsa di questa tecnica per analizzare le credenze e i significati attribuiti dalla popolazione ai concetti della salute e della malattia (Wilkinson, 1998). Come sottolinea Jovchelovitch (2004), le pratiche comunicative quotidiane che sostengono, trasformano e producono le rappresentazioni sociali, sarebbero della stessa identica natura delle pratiche comunicative che hanno modo di svilupparsi in una situazione di dialogo di gruppo quale quella del focus group. «Il focus group può di fatto essere considerato come una società pensante in miniatura» (Kitzinger, Markovà and Kalampalikis, 2004: 239): in questo modo «il metodo di investigazione ha il pregio di riflettere le dinamiche di produzione delle stesse rappresentazioni che intende studiare» (Jovchelovitch, 2004: 246). Anche Jovchelovitch (2004), nello spiegare la particolare utilità del focus group nell'ambito di differenti culture, sottolinea fortemente la necessità di una comprensione dei processi "gruppali" e rappresentazionali che tenga sempre conto della complessità culturale del contesto: le strategie comunicative che si sviluppano nel gruppo, quelle centrali appunto per avere accesso alle rappresentazioni sociali, dipendono in realtà da codici culturali più ampi, ai quali, secondo la ricercatrice, è ugualmente possibile accedere attraverso la metodologia in questione. Il focus group sarebbe dunque lo strumento migliore per entrare in contatto con la diversità culturale, per far emergere, in una situazione di pluralità di gruppo, la logica di produzione e l'eterogeneità delle rappresentazioni sociali.

Nel voler cogliere aspetti attinenti al campo della diversità culturale attraverso l'impiego di questa tecnica, tuttavia, Jovchelovitch (2004) sottolinea come vada tenuto ben presente il fatto che non tutte le culture attribuiscano lo stesso valore alla pratica conversazionale, l'assunto su cui, fondamentalmente, si basa il focus group: il presupposto secondo il quale le persone, in un contesto di gruppo, discuteranno tra loro e si metteranno a parlare liberamente, risulta essere infatti uno dei presupposti fondamentali nelle società urbane e industrializzate occidentali; il dibattito e l'argomentazione, invece, presso alcune comunità tradizionali, sono addirittura giudicati come sconvenevoli. È importantissimo, allora, che il ricercatore cerchi di comprendere come le differenti comunità gestiscano il linguaggio e l'esperienza della vita di gruppo, al fine di adattare, a seconda del contesto, i criteri che il focus group deve possedere per poter accedere a dati interessanti. In caso

contrario, sostiene l'autrice, l'utilizzo di questa tecnica potrebbe rivelarsi addirittura infruttuoso.

3 L'analisi del testo

Molti autori (Corrao, 2000; Zammuner, 2003; Krueger, 1994; Stewart and Shamdasani, 1990) risultano concordi nel giudicare il processo di analisi dei dati del focus group particolarmente difficoltoso, soprattutto a causa della complessità delle informazioni, molte delle quali derivanti da interazioni tra i membri del gruppo. Tale processo di analisi risulta, dunque, non solo oneroso, ma anche peculiare e unico rispetto all'analisi condotta in seguito all'applicazione di altri metodi di ricerca.
Qualunque sia, in ogni caso, il metodo scelto dal ricercatore per analizzare le informazioni ottenute, sostiene Corrao (2000), la base indispensabile per l'analisi di un focus group è sempre costituita dalla trascrizione integrale, previa la sua registrazione, della discussione di gruppo: essa ne consente, infatti, una lettura completa, contenente tutti gli interventi nel dettaglio e nel linguaggio originale in cui sono stati prodotti. Anche tale trascrizione, secondo Ricolfi (1997), consente di aumentare l'ispezionabilità e la ripetibilità, da parte di altri ricercatori, della base empirica della ricerca, aspetto questo di cui la ricerca qualitativa può risultare talvolta deficitaria.
Tra le varie metodologie di cui il ricercatore può servirsi al fine di analizzare i contenuti emersi da un focus group, la descrizione completa delle quali esula dagli obiettivi e dalle caratteristiche della presente trattazione, uno dei metodi interpretativi qualitativi frequentemente impiegati nella ricerca sociale è rappresentato dall'interessante approccio denominato *Grounded Theory* (Glaser e Strauss, 1967; Strauss e Corbin, 1990). La particolarità di tale approccio rispetto ad altri, come verrà spiegato più dettagliatamente in seguito, risiede nella possibilità di far procedere di pari passo l'osservazione e l'elaborazione teorica, in un'interazione continua, circolare e ricorsiva.

4 La ricerca

Per questo studio è stato selezionato un campione di trenta soggetti, tutti i partecipanti svolgevano in Italia al momento del loro reclutamento per lo studio la *professione di infermiere.* L'aspetto cruciale in questo campione è rappresentato tuttavia dalla *provenienza geografica* dei partecipanti

selezionati: l'interesse verso l'approfondimento delle peculiarità culturali attraverso prima un confronto separato e poi un incontro tra differenti culture sui temi della salute/malattia e sulle prassi di cura ha richiesto la costituzione di un gruppo "multi-etnico" di individui, con la presa in considerazione di tre nazionalità. Le tre aree geografiche coinvolte in questo studio sono rappresentate dall'Est Europa, dal Sud America e dall'Italia stessa.
Nonostante i migranti provenienti dal Sud America e dall'Est Europa, in particolar modo le donne, non siano gli unici a trovare impiego in Italia grazie a questo tipo di occupazione, stando a quanto riportato dalle statistiche (Caritas, 2007) questi gruppi etnici sono impiegati in questo settore con una percentuale molto alta, ed è stato questo il motivo che ha portato ad includere nel campione queste due specifiche provenienze geografiche.
La costituzione di un gruppo di infermieri di nazionalità italiana è invece risultato utile non solo come termine di paragone in sede di confronto separato tra i gruppi, ma anche perché ha permesso di mettere alla prova una situazione di scambio e costruzione di un sistema di valori condiviso in riferimento a ciò che è alla base di quest'identità professionale: primariamente i concetti di salute, malattia e delle pratiche di cura.
È stato così possibile suddividere il campione totale composto dai trenta infermieri in tre gruppi sulla base della loro etnia:

- 10 partecipanti italiani
- 10 partecipanti provenienti da un Paese del Sud America
- 10 partecipanti provenienti da un Paese dell'Est Europa.

La procedura di ricerca comprendeva complessivamente *cinque incontri di focus group* con un totale di trenta partecipanti. Tutti i partecipanti svolgevano o avevano svolto in precedenza e per lungo tempo la professione di infermiere. Ciascun soggetto ha offerto la sua collaborazione alla ricerca partecipando a due focus group, il primo assieme ad altri partecipanti della medesima etnia, il secondo assieme a partecipanti di tutte e tre le etnie prese in considerazione. I cinque focus group nell'ordine sono stati organizzati come segue:

- **FASE I:** *Gruppi "intra-etnici".* Sono stati previsti in una prima fase tre incontri di focus group volti a suscitare le rappresentazioni della salute,

della malattia e delle prassi di cura, ognuno con dieci partecipanti provenienti dalla medesima cultura. Gli incontri previsti nella FASE I ("intra-etnica") si sono articolati come segue:

- 1 focus group intra-etnico con 10 partecipanti provenienti dall'area geografica dell'Est-Europa
- 1 focus group intra-etnico con 10 partecipanti provenienti dall'area geografica del Sud America
- 1 focus group intra-etnico con 10 partecipanti italiani

- **FASE II:** *Gruppi "inter-etnici".* In una seconda fase sono stati previsti due incontri di focus group per approfondire le tematiche emerse nei primi tre incontri sui temi della salute, della malattia e sulle prassi di cura. Gli incontri previsti nella della Fase II ("inter-etnica") si articolano come segue:

- 1 focus group multi-etnico con 15 partecipanti di cui:
  - ✓ 5 provenienti dall'Est-Europa
  - ✓ 5 provenienti dal Sud America
  - ✓ 5 italiani
- 1 focus group multi-etnico con 15 partecipanti di cui:
  - ✓ 6 provenienti dall'Est-Europa
  - ✓ 4 provenienti dal Sud America
  - ✓ 5 italiani

Nel presente lavoro di ricerca l'importante fase di creazione, formulazione e stesura delle domande da porre ai partecipanti è avvenuta in *due momenti*: in una fase preliminare di pianificazione della ricerca, chiarito il problema da esaminare e tenendo presente le modalità scelte per la finalità di indagine, descrizione e confronto delle rappresentazioni di salute, malattia e prassi di cura, sono state definite le domande da utilizzare nella Fase I della ricerca; la stessa procedura poi è stata ripetuta non appena ultimati i primi tre incontri, per preparare le domande da proporre nella Fase II della ricerca, i due incontri "multi-etnici". Grazie alla trascrizione integrale dei primi tre focus group e alle riflessioni che sono scaturite da una prima analisi dei contenuti emersi, sono state formulate delle domande corrispondenti ad alcuni temi che potevano essere approfonditi ulteriormente o su cui pareva interessante che i partecipanti si confrontassero.

5 Risultati

Sulla base di quanto emerso nei focus group è stato possibile individuare delle differenze nei trascritti delle discussioni attinenti sia ai contenuti che alla forma. Nello specifico, in relazione alla "salute" i focus monoetnici hanno evidenziato quanto segue:

- Per i migranti provenienti dall'Europa dell'Est, alla salute dell'individuo concorrono una varietà di fattori che possono essere inclusi nei quattro nuclei tematici principali (Psiche, Soma, Individuo, Contesto), pressappoco come accade per le altre due etnie. Si evidenzia, tuttavia, come, sulla base di una struttura generale condivisa tra le tre culture, le differenze tra le rappresentazioni sono costituite in larga misura da "differenze quantitative" riguardo l'intensità con cui un certo tema è stato collegato alla questione centrale (la salute), da differenze riguardo le relazioni tra le parti e da differenze relative a quali temi risultano essere presenti all'interno delle macro-categorie concettuali comuni. Nello specifico, nella rappresentazione della salute delle badanti est-europee, tra i fattori che determinano il benessere dell'individuo, la *sfera psicologica* riveste un'importanza cruciale. Nonostante essi ritengano che il corretto funzionamento dell'organismo, e dunque della *sfera somatica* dell'uomo, abbia effetto e concorra a determinare lo stato di salute, l'importanza di questo fattore risulta appena accennata ("La salute io penso altro modo… nostri organi, nostri organi, questo cuore, fegato… questo penso che è la salute"). È soprattutto sul versante relativo ad aspetti attinenti alla sfera psicologica che invece insistono i discorsi di questo gruppo di partecipanti. Da una parte, infatti, così come nelle altre etnie, alcuni tratti psicologici e temperamentali della persona sono considerati in grado di promuovere lo stato di salute, di avere effetto su di essa; dall'altra parte, invece, vengono sottolineati gli effetti che lo stesso stato di salute dell'individuo può sortire sulla sua sfera psicologica. Tra le "caratteristiche psicologiche" giudicate particolarmente influenti rispetto allo stato di salute all'interno della rappresentazione, la più saliente, risulta sicuramente essere *la forza interiore* dell'individuo. Nonostante non sia presente in maniera esclusiva in questo gruppo culturale, questo concetto rappresenta senza dubbio la peculiarità più evidente delle rappresentazioni delle partecipanti est-europee: esso ricorre in maniera continua nei loro

discorsi tanto rispetto al tema della salute che rispetto ai temi della malattia e della cura, svolgendo dunque un ruolo centrale in ognuna di queste rappresentazioni.

- Anche i partecipanti provenienti dal Sud America considerano la *sfera somatica* e la *sfera psicologica* come due aspetti importanti da cui dipende la salute. Anche per questo gruppo, la psiche è considerata in grado di avere un forte effetto nel determinare la salute ("Quando stai bene… quando stai bene.. tu stesso, credo che dipende la mente, no?") e alcune caratteristiche attinenti al funzionamento psicologico della persona sono indicate come importantissime al fine del mantenimento del benessere: ritorna, seppure con una frequenza inferiore rispetto all'Est-Europa, la necessità dell'ottimismo e di un'*attitudine positiva verso la vita*. Ciò che contraddistingue questo nucleo concettuale nella rappresentazione dei badanti sudamericani è la presenza di un concetto poco definito, ma ripetuto così spesso da non poter essere trascurato, e in ogni caso inseribile nella macro-categoria degli aspetti psicologici: l'idea che la salute dipenda dall'*amore* che la persona ha, e che questo amore sia in qualche modo la base del suo benessere. Si arricchisce e si intensifica il riferimento agli effetti benefici sulla salute di *relazioni interpersonali* sane e positive, che vengono pensate sia all'interno del contesto familiare, dove la salute trae vantaggio da un clima armonioso ("L'armonia familiare pure dipende da salute, perché come dice lui, questa signora no? Che lei è molto, non era malata cioè però… la depressione di tutto quanto di stare sola e… non vedere neanche l'armonia dei figli, si sta ammalando!"), sia in senso più ampio, dove la salute viene ad essere determinata e composta in gran parte anche da un benessere che attiene in generale alla positività delle *relazioni sociali e interpersonali* nelle quali l'individuo è inserito. All'interno della rappresentazione "sudamericana", si potrebbe dire che l'influsso sulla salute di componenti riferibili al nucleo tematico del *contesto/ambiente* in cui si colloca idealmente l'individuo venga sottolineato in maniera più marcata rispetto alle rappresentazioni degli altri gruppi.

Ciò che, in ultima analisi, risulta interessante notare in questa rappresentazione dei migranti sudamericani, è la linea di connessione è possibile tracciare tra ciò che attiene all'"Individuo" e ciò che attiene al "Contesto/Ambiente" in relazione alla salute. È l'idea di

un'interazione tra queste due componenti: il modo in cui l'individuo interagisce con l'ambiente che lo circonda, le sue modalità di reagire agli innumerevoli stimoli che esso pone, si va conseguentemente a riflettere sul proprio stato di benessere ("Essere in salute comporta la capacità di adattarsi fronte agli stimoli esterni e agire ad essi in maniera positiva"). L'esistenza di questo "ponte" fondamentale tra individuo e ambiente viene accennata solamente nell'ambito di quest'etnia.

- È stato piuttosto singolare, nell'andare a comporre la mappa dei concetti messi in collegamento con il tema della salute dai partecipanti italiani, imbattersi in una quantità nettamente minore di concetti rispetto agli altri due gruppi, con una conseguente minore possibilità di dedurre le relazioni intercorrenti tra questi e il tema principale, all'interno della Rappresentazione.

Il nucleo tematico relativo al *contesto/ambiente*, risulta essere completamente assente, in quanto nel corso della discussione non è emerso alcun riferimento a questi aspetti, mentre tutte le altre aree tematiche sono risultate presenti, seppur solamente accennate e poco approfondite. Si potrebbe ipotizzare, in accordo con quanto riportato da Herzlich (1973), che la scarsità delle argomentazioni direttamente riferite al tema della salute potrebbe avere, in questo gruppo culturale, il significato forse di un'accentuata tendenza a sconfinare, parlando di salute, inevitabilmente nell'ambito di quella che è considerabile come la vera e propria altra faccia della medaglia, il suo "opposto": la malattia. In questo gruppo, infatti, ad una rappresentazione più scarna della Salute, ne corrisponde una molto variegata e dettagliata della Malattia.

Considerata l'ipotesi secondo la quale, sulla base del forte influsso che i gruppi di discussione eserciterebbero sul cambiamento delle rappresentazioni individuali (Echebarria Echabe and Gonzalez Castro, 1999), nei focus group multi-etnici ci si sarebbe dovuto attendere una qualche modificazione nelle idee sulla salute, sulla malattia e sulla cura rispetto a come precedentemente esposte separatamente da ogni gruppo. In sede di analisi dei trascritti si è cercato di prendere in considerazione questo aspetto, rintracciando quali cambiamenti fossero avvenuti nel passaggio dalla prima alle seconda tipologia di discussione di gruppo.

Rispetto al modo in cui essi erano emersi nella fase di confronto separato tra i gruppi, durante i focus group multi-etnici i partecipanti si sono confrontati con domande attinenti soprattutto ai principali aspetti di dissimilarità osservati sulla base di una prima lettura del materiale a disposizione. Premeva, infatti, dare luogo ad un confronto tra i partecipanti che, più che sull'ampia base comune riscontrata nelle loro rappresentazioni della salute si concentrasse piuttosto sulle questioni in merito alle quali si erano osservate, nei diversi gruppi, delle concezioni non univoche, favorendo la discussione su di esse e ponendo particolare attenzione affinché tutti esprimessero la propria opinione in merito.

Sommando gli elementi emersi da ambedue le esperienze di discussione "multi-culturale", in effetti, è stato possibile osservare come, all'interno delle tematiche, il principale cambiamento riscontrabile fosse rappresentato da una sorta di "normalizzazione" di alcuni aspetti peculiari delle rappresentazioni di ogni etnia.

Questa condizione si è potuta osservare principalmente considerando quali frequenze fossero assunte dai diversi codici, nei focus group inter-etnici: le frequenze di concetti che, nei focus group mono-etnici erano risultati caratteristici in una certa etnia, piuttosto che nelle altre, appaiono così livellarsi quando i partecipanti appartenenti ai tre gruppi hanno la possibilità di discuterne insieme. Questo non solo per un minore apporto di contributi da parte dei partecipanti di quell'etnia in relazione a quei concetti, ma anche per la loro uguale distribuzione di frequenza nei tre gruppi etnici, contrariamente a quanto era accaduto in sede di confronto separato. Come se, attraverso il contatto tra persone appartenenti a differenti culture, nel tentativo di far riferimento ad una base comune, le differenze si attenuassero, e ogni partecipante riuscisse a trovare nella propria esperienza i codici per decifrare e condividere ciò che viene sostenuto dall'altro.

Riferimenti bibliografici

Aranda, M.P. and B.G. Knight (1997) 'The influence of ethnicity and culture on the caregiving and coping process: A socio-cultural review and analysis'. *The Gerontologist*, 37, pp. 342-54.

Bakker, W., J.P. van Oudenhoven and K.I. van der Zee (2004) 'Attachment styles, personality, and dutch emigrants intercultural adjustment'. *European Journal of Personality*, 18(5), pp. 387-404.

Berry, J.W. (1997) 'Immigration, acculturation e adaptation'. *Applied Psychology: An International Review*, 46(1), pp. 5-34.
– (2008) 'Globalizzation and acculturation'. *International Journal of Intercultural Relations*, 32, pp. 328-36.
Caritas (2007) *Immigrazione. Dossier statistico 2007*. Roma: Idos.
Corrao, S. (2000) *Il focus group*, Milano: Franco Angeli.
Echebarria Echabe, A. and J.L. Gonzalez Castro (1999) 'Group Discussions and Changes in Attitudes and Representations'. *The Journal of Social Psychology*, 139 (1), pp. 29-43.
Flick, U. (1998) *An Introduction of qualitative research*. London: Sage.
Fortier, J.P. and D. Bishop (2003) Setting the agenda for research on cultural competence in health care, in C. Brach (ed.) *Final report: Setting the agenda for research on cultural competence in health care*, Washington, DC: U.S. Department of Health and Human Services, pp. 11-4.
Glaser B.G. and A.L. Strauss (1967) *The discovery of grounded theory: strategies for qualitative research*. Chicago: Aldine.
Herzlich, C. and D. Graham (1973) *Health and Illness. A social Psychological Analysis*. London: Academic Press Inc.
Janevic M.R. and C.M. Connell (2001) 'Racial, ethnic, and cultural differences in the dementia caregiving experience: Recent findings'. *The Gerontologist*, 41, 334-47.
Jovchelovitch, S. (2004) 'Contextualiser les focus groups: comprendre les groups et les cultures dans la recherche sur les representations'. *Bulletin de Psychologie*, 57(3), 471, pp. 245-252.
Kitzinger, J., I. Markovà, and N. Kalampalikis (2004) 'Qu'est-ce que les focus groups?'. *Bulletin de Psychologie*, 57(3), 471, pp. 237-43.
Krueger, R.A. (1994) *Focus Group. A Pratical guide for Applied Research*. London: Sage.
Kurman, J. and C. Ronen-Eilon (2004) 'Lack of knowledge of a culture's social axioms and adaptation difficulties among immigrants'. *Journal of Cross-Cultural Psychology*, 35(2), pp. 192-208.
La Framboise, T., H.L.K. Coleman and J. Gerton (1993) 'Psychological impact of biculturalism: Evidence and theory'. *Psychological Bulletin*, 114(3), pp. 395-412.
Leventhal, H. T.J. Musumeci and R.J. Contrada (2007) 'Current issues and new directions in Psychology and Health: Theory, translation,

and evidence-based practice'. *Psychology and Health*, 22, pp. 381-5.
Palumbo, M. (2001) *Il processo di valutazione. Decidere, programmare, valutare*. Milano: Franco Angeli.
Portes, P.R., M.F. Zady and R.M. Dunham (1998) 'The effects of parents' assistance on middle school students' problem solving and achievement'. *Journal of Genetic Psychology*, 159 (2), pp. 163-78.
Ricolfi, L. (1997) *La ricerca qualitativa*. Roma: La Nuova Italia Scientifica.
Sondik, E., J. Lucas, J. Madans and S. Smith (2000) 'Race/ethnicity and the 2000 census: implications for public health'. *American Journal of Public Health*, 90, pp. 1709-13.
Stagi, L. (2000) 'Il Focus Group come tecnica di valutazione. Pregi, difetti, potenzialità'. *Rassegna Italiana di Valutazione*, 20, pp. 67-88.
Stewart, D.W. and P.N. Shamdasani (1990) *Focus Group. Theory and Practice*. London: Sage.
Strauss, J. and A. Corbin (1990) *Basics of Qualitative Research. Grounded Theory Procedures and Techniques*, Newbury Park: Sage.
Valle, R. (1998) *Caregiving across cultures: Working with dementing illness and ethnically diverse populations*. Washington, DC: Taylor & Francis.
Velotti, P. and G.C. Zavattini (2008) 'Le professioni d'aiuto nelle società multietniche: differenze e similitudini nelle rappresentazioni della salute e della malattia'. *Ciclo Evolutivo e Disabilità*, 11, 1, pp. 101-11.
Ward, C. and I. Styles (2003) 'Lost and found: Reinvention of the self following migration'. *International Journal of Applied Psychoanalytic Studies*, 5(3), pp. 349-66.
Wilkinson, D. (1995) Central civilization, in S.K. Sanderson (ed.) *Civilization and world systems*, London: Altamira Press, pp. 198-9.
Zammuner, V.L. (2003) *I focus group*, Bologna: il Mulino.

Parlare italiano: towards a multidimensional description and a multidisciplinary explanation

Miriam Voghera
University of Salerno
voghera@unisa.it

Abstract

According to numerous research works, there exist some constants of spoken language that cannot be referable only to the use of the phonic-auditory channel. In fact, a correct explanation of speech structures must necessarily take into consideration the semiotic and cognitive potentials and constraints that are normally associated to speaking and listening. Only in this way would it be possible to construct a general form of grammar capable of explaining speech within a unitary framework and not as an eccentric and/or marked appendix. Assuming this point-of-view requires linguistic analysis crossing over the traditional linguistic partition boundaries and the concurrence of knowledge that derives not only from linguistics, but also from other disciplines such as philosophy, psychology, anthropology, etc. A piece in the puzzle of this long-term programme of general interest is the constitution of a permanent observatory of speech which, while continuing to produce new material, inserts the data into a common framework. To this end a joint project, promoted by different Italian research groups, was launched in 2004: Parlare italiano: osservatorio degli usi linguistici. *All the material and results obtained during the project to date may be consulted online at the portal www.parlaritaliano.it, which represents the structure and internal organization of the entire project.*

1 Introduction

Thanks to considerations made during the last few years and the enormous descriptive efforts which were also possible thanks to the collection of text corpora and their deriving analysis (De Mauro *et al*, 1993; De Mauro, 1994; Biber *et al*, 1999; Albano Leoni, 2003; Cresti and Moneglia, 2005; Albano Leoni and Giordano, 2005; Cutugno and Savy, in press), we presently believe that we have a relatively clear picture of the principal structures of the spoken language. In fact, not only we have a better knowledge of most typical and frequent spoken

uses, but we are especially more conscious of the mechanisms of signification that underlie single linguistic choices.

Following an important investigation period principally with documentary objectives (De Mauro, 1994; Biber, 1988; 1995, whose primary cognitive aim was to bring to light less known (and sometimes even unknown) language phenomena and uses, a new stage is possible, aimed for the construction of a systematic framework, within which both speech phenomena and spoken structures can be explained. In order to reach this objective, it is necessary not only to keep in mind the type and frequency of various language structures, but to find their reason and specificity in relation to all the elements contributing to their meaning. In fact, the presence of determined language structures within a given text not only depends on what is allowed by grammar and speaker's individual choices, but also by what the enunciating condition render more functional and preferable. As has already been stated in other research studies (Voghera, 2008; and bibliography within), there exist some *constants* of spoken language that cannot be referable only to the use of the phonic-auditory channel, but rather depend on the complex relation between language grammar and the semiotic and cognitive conditions in which the language structures take place. In fact, a correct explanation of speech structures cannot be limited to appealing to the phonic aspects, but must necessarily take into consideration the semiotic and cognitive potentials and constraints that are normally associated to speaking and listening. When we speak (and listen), perhaps even more than when we use other modalities of communication, the pressure- so to speak - of enunciating constraints is evident. In fact, it proves to be decisive in explaining the spoken usages, the interdependence between language structures, production/reception conditions and relation setting between participants during the communicating act.

This makes speech a privileged observation point for the study of relations between language structures and signification conditions. Including speech among the data which we build our theories, therefore broadens not only knowledge regarding the variability of uses, but also that of the linguistic system, of its components or levels; this might lead to the construction of models that have a higher degree of explanatory adequacy.

Even the analysis of minute facts must therefore be inserted within a broader framework that keeps the general semiological context in

mind, along with the inter-active dimension within which spoken usage finds its (precisely) significant reasons. Only in this way would it be possible to construct a general form of grammar capable of explaining speech within a unitary framework and not as an eccentric and/or marked appendix. In our opinion, this is a necessary step that must be taken in order to single out the basic and fundamental elements of speech (but even of other modalities of communication) without ever losing touch with the constituent uniqueness of language.
On one side, assuming this point-of-view requires linguistic analysis crossing over the traditional linguistic partition boundaries as it compares phenomena of different levels that are apparently not immediately connected to one another- one might consider, for instance, the pertinence of phonetic properties of the super-segmental level for the syntactic articulation of speech; and, on the other side, the concurrence of knowledge that derives not only from linguistics, but also from other disciplines such as philosophy, psychology, anthropology, etc. In other words, a good representation of what occurs when we speak and listen presumes a multidimensional description and a multidisciplinary explanation.

2 An observatory of spoken Italian

A piece in the puzzle of this long-term programme of general interest is the constitution of a permanent observatory of speech which, while continuing to produce new material, inserts the data into a common framework. To this end a joint project, promoted by different Italian research groups, was launched in 2004: *Parlare italiano: osservatorio degli usi linguistici.* All the material and results obtained during the project to date may be consulted online at the portal www.parlaritaliano.it, which represents the structure and internal organization of the entire project.
The observatory has two strongly related objectives. The first objective is that of broadening the cognitive basis of enunciative and grammatical mechanisms of spoken communication through research based on corpora. The second objective entails evaluating how and how much the broadening of linguistic data might contribute to a better understanding of the linguistic system in its entirety. In fact, it is evident that speech, at an initial level, is characterized by a sub-group of typical language structures that are not (or only partially)

observable in other contexts; at a second level, speech allows us to discover the relation between different portions of grammar, otherwise hidden, but nonetheless central to the general architecture of the system.
Computational objectives addressed to researching the structure of metadata, linguistic databases, analysis and implementation are added to the strictly linguistic objectives. The project includes research on theoretical and practical implementation of techniques for automatic segmentation of audio and video signals, with specific attention to their use in the management of multimedia and multimodal linguistic corpora is carried out within the project.
Considering these objectives, *Parlare italiano* gathers together numerous researchers from different research groups who are involved in the many aspects of spoken communication. This is a driving force that allows not only to present a broad panorama of studies and points-of-view, but especially to create the necessary framework for the construction of a systematic vision of linguistic facts, being a priority when setting apparent eccentricities in order.
The multiplicity of approaches can be derived from the research fields present within the project, which cover many aspects of research regarding speech and spoken communication: from the history of studies on speech, to research regarding diachronic, regional, and diafasic registers as well as difference between native and non-native speech, form the contributions of computational linguistics to experiments in the technological field for the analysis and management of speech data, to the contributions of audiological and psychological research regarding speech disorders. Although already quite wide, the list may not be considered exhaustive or concluded, but open to amplification and enrichment.
Many description levels are also represented: prosodic, morphologic, syntactic, lexical, semantic and pragmatic.

2. 1 Corpora, tools and corpus-based investigations

As has already been mentioned in the introduction, the collection and analysis of the corpora has constituted an important turning-point for speech studies. Although the amount of written corpora is still very much bulkier, various speech corpora in different languages are now available. Nevertheless, corpora are not always freely accessible to

scientific communities substantially due to reasons of an editorial character.
Therefore, the choice of *Parlare italiano* has been to offer utmost publicity to all the stages of the projects that have been presented, not only supplying the results but also publishing all the various procedures that have led to the various scientific conclusions in the most transparent manner possible. For this reason, in addition to the corpora collected, we have also published the tools used for the annotations and analyses so as to warrant their use during further research studies conducted by other researchers. This choice offers external subjects the chance to concretely follow our theoretical and applicative options.
At present, *Parlare italiano* has published nine new corpora, aimed for investigations into different linguistic areas. We have listed below the authors and some summarized details for each of these.

AN.ANA.S. (M. Voghera, G. Turco, A. Landolfi): treebank made up of syntactic tagging of Italian spoken (Corpus Clips) and written (Corpus Penelope) texts (Voghera and Turco, 2008; Voghera and Cutugno, in press). Audio files and orthographic transcriptions of the spoken material can be freely downloaded from the portal www.parlaritalino.it as well as the editing and management software, Xgate, developed in the field of the project (Cutugno and D'Anna, 2008, in press).

AN.ANA.S. L2 (G. Turco, M. Voghera): Treebank of the syntactic tagging of written texts produced by students learning Italian as a second language (Turco and Voghera, in press). The searchable database generated by the treebank can be consulted using the management software Xgate, developed in the field of the project, freely downloadable from the portal www.parlaritaliano.it.

European languages PraTiD Corpus (R. Savy, S. De Leo, I. Alfano, I. Solis Garcia, E. Cioffi): made up of *task-oriente*d dialogues in English, German, Spanish, Portuguese and French. These have been annotated according to a pragmatic annotation scheme (PraTiD) that envisages a hierarchical structure of the analysis categories (Savy and Solis, 2008). The architecture of annotation is in the XML format, freely downloadable from the portal www.parlaritaliano.it.

Legibility and lexicon of the political-parliamentary language (E. Piemontese): made up of texts from the 14th and 15th legislature of the Italian Senate; the texts have undergone legibility analysis using the Èulogos-CENSOR formula.

LIPS (M. Vedovelli, A. Pallassini, S. Machetti, M. Barni, C. Bagna, S. Pieroni, F. Gallina): made up of approximately 100 hours of speech by learners of Italian recorded during the *CILS - Certificazione di Italiano come Lingua Straniera* examinations at the University for Foreigners in Siena (Vedovelli, 2006). In addition to the orthographic transcription of the corpus, the deriving frequency lexicon is also available (Barni and Gallina, 2008).

Movie language (A. Giannini, M. Pettorino, I. Vitagliano): made up of four movies filmed in different years and periods during a 65-year time-span (Vitagliano, Giannini and Pettorino, 2009). The corpus is provided with: a) audio recordings in .wav format of the selected sequences, subjected to acoustic spectrum analysis and tagging; b) orthograpfhic transcriptions of the analyzed *corpus*, carried out both on a perceptive and an acoustic basis; c) comparison between the data analyzed, gained from the analysis of the corpus.

Television news speech: made up of RAI (national broadcast company) (A. Giannini, M. Pettorino, I. Vitagliano): television news recordings, complete with orthographic transcriptions on an acoustic basis and an initial elaboration of the data, for a comparison between speech during the Sixties and contemporary speech (Pettorino and Giannini, 2005).

Pra.Tid. (R. Savy, S. De Leo, M. Castagneto): made up of *task-oriente*d dialogues, recorded according to a pragmatic annotation scheme (Pra.TiD) that envisages a hierarchical structure of the analysis categories (Savy and Castagneto, in press). The architecture of annotation is in the XML format, freely downloadable from the portal www.parlaritaliano.it.

SpIt-MDb (*Spoken Italian – Multilevel Database*) (R. Savy, C. Crocco, G. Turco, F. Cutugno, G. Clemente, C. Napoletano): made

up of dialogues annotated at many levels and in different formats, converted into a *database* that can be consulted through the application of the AG-SpIt *toolkit* (based upon AGTK techniques by Bird and Liberman, 2001), freely available at the portal www.parlaritaliano.it.

In addition to new corpora, *Parlare italiano* published corpora previously collected, but difficult to consult either because out of print or never entirely published:

AVIP-API (*Archives of Italian Speech* Crocco, Savy, Cutugno, 2003): made up of a) *map-task* dialogues collected in Naples, Pisa, Bari, Brindisi and Florence; b) lists of words (toponyms of maps) read by those who participated in the dialogues; c) children's speech (normal hearing and deaf children); d) sentences read and collected in Pisa, for the in-depth study of specific phenomena of post-lexical nature.

Penelope (Policarpi and Rombi, 2005): made up of Italian texts, both written and oral, ancient and contemporary, with very narrow sampling – small collections for few periods by text or by author – trying to cover the highest possible number of language registers. The corpus can be searched on the basis of the text type.

Some of the listed corpora are of small dimensions: in fact we have been guided by the idea, very well expressed by Sampson (2003) in his "small is beautiful", according to which accuracy and explicitness of the annotation is often more important than quantity.
In addition to the corpora already enumerated, *Parlare italiano* has published other works based upon these or upon other material. Again in this case we have on disposal both the linguistic data and the material explaining the procedures that have been used. Amongst others, we wish to point out work on:

- phonetics, regarding the perception of accents in Italian and Spanish (Alfano, Llisterri and Savy, 2007);
- derivational morphology carried out on the LIP Corpus (Iacobini and Adinolfi, 2008);
- inflectional morphology carried out upon the Italian section of the Corpus C-Oral Rom (Koesters Gensini and Aureli);

- semantic weight carried out on the Corpus AN.ANA.S. and the lexicographical treatment of spoken usages in the "Grande Dizionario italiano dell'uso" (Basile and Diodato, in press);
- lexicon, regarding the use and the frequency of exotisms in parliamentary lexicon (Villani).

Further research material on various aspects of spoken language is available within the manuscript and publication sections.

2. 2 Interlinguistic comparisons

Despite the fact that the observatory is principally dedicated to the Italian language, during the last few years a research area dedicated to interlinguistic comparison has been developed. One of the objectives of the project is actually investigating into the more general and interlinguistically shared spoken features, paying particular attention to the interactive dimension of speech and to the inter-subjective and inter-cultural differences observable in the various languages through the adoption of different styles of communication.

To this end, we have begun to collect a corpus of speech in various European languages: French, English, Spanish, Portuguese and German (cfr. § 2.1 European languages PraTiD Corpus). The corpus, made up of *task-oriented* dialogues, is available both in the audio file version and transcription. Such a corpus constitutes the ideal basis for comparing grammar structures and linguistic use between languages or between language varieties because it both offers the possibility to delve into homogeneous data and extrapolate comparable information in a simple and direct manner and also puts pre-analyzed data at disposal.

Other research works on interlinguistic comparisons have been undertaken in syntactic analysis. In addition to the annotation of a corpus of both spoken and written Italian, annotation has begun for spoken English and Spanish texts through the use of the already mentioned system for syntactic annotation AN.ANA.S. (Voghera and Cutugno, in press). The system generates a tagged text on the basis of XML standards and *Text Encoding Initiative* (TEI), finally obtaining a database that can consulted for the acquisition of quantitative and qualitative data regarding the syntactic characteristics of the corpora that have been analyzed (cfr. § 2.1).

Particularly developed is the research dedicated to the comparison between native and non-native Italian. *Parlare italiano* has published the largest corpus of Italian L2, LIPS (cfr. § 2.1), and host a project regarding the acquisition of discourse markers in the Italian spoken by Spanish learners, promoted by Carla Bazzanella together with Spanish scholars and schools teachers.

3 Linguistic theories and applications.

Parlare italiano aims at also becoming a reference point for the development of applications. Obviously the application environments are many: we wish to only mention those two that we believe are of particular interest and usefulness in the field of speech studies.

An initial field of application regards educational linguistics. The construction of a general grammar is crucial for the creation of educational applications aimed at the integrated development of all linguistic abilities. Despite the fact that language education, especially second language education, has tried to develop listening and speaking abilities, there is still a lack of consolidated general reference models, especially for the teaching of the mother tongue. This produces contradictory and paradoxical effects. On one hand speech is, so to say, taken for granted: a primary and natural ability that is not learned and therefore is not taught. On the other hand, the spoken language, since it is more informal and less controlled than the written language, is strongly censored and "corrected". This is all the moreso for teaching Italian as a second or foreign language, in which spoken features are often mistaken for elements of incompetence (Ambroso and Bonvino, 2008).

There is a lack of awareness of the fact that even speech, just like the written language, presents an great range of registers that one must teach/learn to use, according to context variation. Hence it is necessary to design an itinerary of linguistic education that should not be less systematic and continuous than that one developed to teach different written registers. In order to learn how to speak in an appropriate and effective manner, it is not only necessary to be exposed to a broad range of verbal stimuli, but it is also necessary to be able to acquire their function within a systematic framework of linguistic, contextual and cultural elements. In conclusion a good linguistic education must guide the student to observe the most

typically spoken structures and compare them with those most typically present in other forms of communication. This is the only way in which the learners may discover how different language uses may be projected on the different portions of the linguistic space. In other words, we believe it is necessary for language education to ensure that speech education be an integral part of the overall linguistic development of the learners.

However, this has not yet become a common and consolidated legacy within pedagogical theories and practices. The result is that there is either little education of speech and listening or it is done in an impressionistic and episodic manner.

There are obviously many roots for this neglect for speech. In the first place, most of the linguistic models, and consequently pedagogical grammars, are prevalently shaped upon the written language. In the second place, until present speech researchers have paid little attention to the elaboration of general reference models applicable on the educational level. It is therefore necessary that theoretical linguistics and educational linguistics work together as one towards this objective.

Parlare italiano offers many cues in this direction, providing informational data on the acquisitional processes of the spoken language for learners of a second language (cfr. § 2.2 Corpus LIPS) and analysis instruments for second languages (cfr. AN.ANA.S. L2, cfr. § 2.2).

A second area of application, which we believe is of great interest, regards corpus linguistics and computational linguistics. The collection of corpora is often accompanied by the development of management and data analysis systems that are more and more sophisticated, through the creation of annotation systems belonging to different levels of linguistic description and quantitative-qualitative databases. The annotation and tagging systems, in almost all linguistic areas, have become very analytical and in many cases allow one to obtain very detailed quantitative and qualitative data (see for instance the different Treebanks which are available today, Abeillé, 2003). These systems are useful and precious for refined linguistic analysis but they do not always communicate with one another. This means that each corpus becomes a product that can be benefited from only singularly.

On the contrary, it is necessary to develop lighter management and analysis systems, perhaps even less sophisticated ones, but that isolate a limited number of basic characteristics of the analyzed material and render the same immediately comparable to other material. It would therefore be useful to create tools that offer the opportunity to build (in a reasonable lapse of time) a sort of identity card of the corpora under different profiles: phonetic-phonological, morphosyntactic, semantic, pragmatic, etc. Of course this should not substitute detailed linguistic analysis, but might be in some way preliminary to them. In other words, it would be desirable to develop some basic measures, so to speak, that would allow to sketch out the general features of the language of different corpora: a sort of corpora identity card. This should integrate sociolinguistic categorizations that does not always select language samples that are necessarily linguistically homogeneous at all levels. Obtaining basic linguistic data and information for each corpus would increase its usability and comparability even by non-linguistics, certainly for teaching uses, but even for different applications and elaborations: for example, in the speaker and text recognition field or *data-mining* development.

The objective is certainly an ambitious one and calls for close collaboration between theoretical and computational linguists and computer scientists – first and foremost on the theoretical level: good application results can only be achieved with good theoretical research work in the background. In fact, it is evident that the individuation of internal analysis criteria depends upon our own capacity to single out the boundary between invariance and possible variations, namely to construct a general language grammar in relation to diamesic, but also diafasic variation. This can only take place through a good meta-linguistic modelling of the data description categories, what computer scientists define as metadata.

Again in this field *Parlare italiano* offers some cues, thanks to the "diagnostical" potential of the many investigations, including research carried out on rhythm by Giannini and Pettorino, on lexicon by the group cordinated by Vedovelli, on syntax by the group coordinated by Voghera and pragmatic analysis by the groups coordinated by Bazzanella and Savy.

In conclusion, the experience which has been gained during the common work carried out during these past few years has allowed us not only to consolidate and broaden the scientific knowledge acquired,

but to picture numerous new lines of theoretical and applied research. The ambition of *Parlare italiano* is to offer a place where these developments can take place, at the same time hosting proposals and research from a greater number of scholars. The first colloquium of *Parlare italiano* (Giornate di Parlare italiano) on the *Intensity phenomena in spoken Italian* (Bazzanella and Gili Fivela, 2009), organized by Carla Bazzanella and Barbara Gili Fivela at the University of Lecce represented a profitable occasion towards this direction.

Acknowledgements

Parlare italiano has been funded by MIUR as a Prin project in 2004 and again in 2006.
Special and sincere thanks go to our colleagues who, along with the national coordination group from the Salerno University, presently constitute *Parlare italiano*: the group from the University of Naples "L'Orientale", coordinated by Antonella Giannini; the group from the University of Turin, coordinated by Carla Bazzanella; the group from the University for Foreigners in Siena, coordinated by Massimo Vedovelli.
My personal gratitude goes to the members of the coordination group: Renata Savy, who tirelessly contributed to the entire construction of the project in all of its stages; Grazia Basile and Claudio Iacobini. The coordination group has also benefited from the precious collaboration of Francesco Cutugno from the NLP-Group at Department of Physics at the University of Naples "Federico II", who, in addition to having planned the portal www.parlaritaliano.it, supervises the development and growth of the computational side of the project.
Many strengths should join together as one in order to coordinate a project that include so many research scientists; this vital strength has been offered by members of the portal's editorial staff: Rosa Giordano, Leandro D'Anna, Simona De Leo and Iolanda Alfano.
We ultimately wish to thank all our colleagues who, over these last years, have contributed in the various stages towards the establishment of *Parlare italiano* and whom we list below in simple alphabetical order:

Federico Albano Leoni, Gennaro Auletta, Massimo Aureli, Carla Bagna, Monica Barni, Cristina Bosco, Ludovica Bracci Laudiero, Fiammetta Carloni, Laura Carrabba, Marina Castagneto, Michele Cavaliere, Michela Cennamo, Giuliana Clemente, Anna Corazza, Emanuela Cresti, Claudia Crocco, Paolo D'Achille, Tullio De Mauro, Sabine E. Koesters Gensini, M. Emanuela Piemontese, Marco Fabbri, Francesca Gallina, Dalia Gamal, Alessandro Garcea, Pasquale Giannini, Barbara Gili Fivela, Fabrizia Giuliani, Ludovica La Bruna, Emilia Landolfi, Carla Laria, Joaquim Llisterri, Edoardo Lombardi Vallauri, Francesca M. Dovetto, Sabrina Machetti, Massimo Maggini, Elio Marciano, Olga Maria Manfrellotti, Paola Micheli, Massimo Moneglia, Adele Munaretto, Carlota Nicolas Martinez, Franca Orletti, Alessandro Pallassini, Alessandro Panunzi, Ester Paone, Massimo Pettorino, Simone Pieroni, Irma Pirri, Elena Pistolesi, Pasquale Riccardi, Giuseppina Santoru, Antonietta Scarano, Manuela Senza Peluso, Raffaele Simone, Laura Sprugnoli, Antonella Stefinlongo, Francesca Tini Brunozzi, Ida Tucci, Giuseppina Turco, Flavia Ursini, Marianna Vallone, Paola Villani, Ilaria Vitagliano.

References

Abeillé, A.L. (ed.) (2003) *Treebank.*: *Building and Using Parsed Corpora.* Dordrecht/Boston/London: Kluwer Academic Publisher.

Albano Leoni, F. (2003) Tre progetti per l'italiano parlato, in N. Maraschio and T. Poggi Salani (eds) *Italia linguistica anno Mille. Italia linguistica anno Duemila*, Bulzoni: Roma, pp. 675-83.

Albano Leoni, F. and R. Giordano (eds) (2005) *Italiano parlato. Analisi di un dialogo*, Napoli: Liguori.

Alfano, R., J. Llisterri and R. Savy (2007) The Perception of Italian and Spanish Lexical Stress: A first cross-linguistic study, in *Proceedings ICPhS* XVI, Saarbrucken, pp. 1793-6.

Ambroso, S. and E. Bonvino (2008) 'Livelli diversi di competenza nella gestione dell'italiano L2. Ipotesi dall'analisi di un corpus'. *Testi e Linguaggi*, 2, pp. 37-65.

Barni M., Gallina F. (2008) Le parole degli stranieri: il LIPS, il primo lessico di frequenza dell'italiano parlato dagli stranieri, in M. Barni, C. Bagna and D. Troncarelli (eds) *Lessico e apprendimenti*, Milano: Franco Angeli, pp. 143-52.

Basile, G. and F. Diodato (in press) Parlato e dizionari. Il trattamento degli usi parlati nel GRADIT, *Proceedings of the International Conference La comunicazione parlata*, Napoli, 23-25 febbraio 2009.

Bazzanella C. and B. Gili Fivela (eds) (2009) *Fenomeni di intensità nell'italiano parlato*. Firenze: Cesati.

Biber, D. (1988) *Variation across speech and writing*. Cambridge: Cambridge University Press.

– (1995) *Dimensions of register variation: A cross-linguistic comparison*. Cambridge: Cambridge University Press.

Biber, D., S. Johansson, G. Leech, S. Conrad and E. Finegan (1999) *Longman Grammar of Spoken and Written English*. Harlow, Essex: Pearson Education.

Bird, S. and M. Liberman (2001) 'A formal framework for linguistic annotation'. *Speech Communication*, 33, pp.23-60.

Cresti, E. and M. Moneglia (eds) (2005) *C-ORAL-ROM Integrated Reference Corpora for Spoken Romance Languages.* Amsterdam: Benjamins.

Crocco, C., R. Savy and F. Cutugno (eds) (2003) *API: Archivio del Parlato Italiano*, CIRASS - Università degli Studi di Napoli Federico II, Napoli, DVD.

Cutugno, F. and L. D'Anna (2008) XGate and XRG: tools for visually editing, querying and benchmarking XML linguistic annotations, in *Proceedings LangTech*, pp. 86-89.

– (in press) Limiti e complessità del recupero dell'informazione da *treebank* sintattiche, in *Atti del XL Congresso SLI*, Vercelli 21-23 Settembre 2006, Roma: Bulzoni.

De Mauro, T. (ed.) (1994) *Come parlano gli italiani*, Firenze: La Nuova Italia.

De Mauro, T., F. Mancini, M. Vedovelli and M. Voghera (1993) *Lessico di frequenza dell'italiano parlato*. Milano: Etas.

Iacobini C. and Adinolfi A. (2008) La derivazione nel parlato dell'italiano messa a confronto con quella delllo scritto, in M. Pettorino, A. Giannini, M. Vallone and R. Savy (eds) *La comunicazione parlata*, Tomo III, Napoli: Liguori, pp. 494-512.

Pettorino, M. and A. Giannini (2005) Italian TV speech: a diachronic analysis, in *Proceedings of IXth International Symposium on Social Communication*, Actas-II: Santiago de Cuba, pp. 849-53.

Policarpi, G. and M. Rombi (2005) Tendenze nella sintassi dell'italiano contemporaneo, in T. De Mauro and I. Chiari (eds) *Parole e numeri. Analisi quantitative dei fatti di lingua*, Roma: Aracne, pp. 139-56.

Sampson, G. (2003) Thoughts on two decades of drawing trees, in A. Abeillé (ed.) *Treebank. Building and Using Parsed Corpora.* Dordrecht/Boston/London: Kluwer Academic Publisher, pp. 23-41.

Savy, R. and M. Castagneto (in press) Funzioni comunicative e categorie d'analisi pragmatica: dal testo dialogico allo schema xml e viceversa, in *Linguistica e modelli tecnologici di ricerca*, Atti del XL Congresso SLI, Bulzoni: Roma.

Savy, R. and F. Cutugno (in press) Clips: diatopic, diamesic, diafasic variations in spoken Italian, in *Proceedings of the fifth Corpus Linguistics Conference*, Liverpool 20-23 July 2009.

Savy, R., F. Cutugno and C. Crocco (2006) Multilevel corpus analysis: generating and querying an AGset of spoken Italian (SpIt-MDb), in *Proceedings of 5th international LREC 2006 Language Resources and Evolution*, Genova, pp. 1654-9.

Savy, R. and I. Solís García (2008) Strategie pragmatiche in italiano e spagnolo a confronto: una prima analisi su corpus. *Testi e linguaggi*, 2, pp. 214-42.

Turco, G. and M.. Voghera (in press) From text to lexicon: the annotation of pre-target structures in an Italian learner corpus, in *Proceedings of the Third International Lablita Workshop in Corpus Linguistics*, Firenze: Firenze University Press.

Vedovelli M. (2006) Il LIPS - Lessico di frequenza dell'italiano parlato dagli stranieri, in C. Bardel and J. Nystedt (eds) *Progetto Dizionario Italiano Svedese* in *Acta Universitatis Stockholmiensis*, Stockholm: Romanica Stockholmiensia, 22, pp. 55-78.

Vitagliano, M., A. Giannini and M. Pettorino (2009) Le emozioni nel tempo, in C. Bazzanella and B. Gili Fivela (eds) (2009) *Fenomeni di intensità nell'italiano parlato,* Firenze: Cesati, pp. 165-78.

Voghera, M. (2008) Progettare la grammatica del parlato, in M. Pettorino, A. Giannini, M. Vallone and R. Savy (eds) *La comunicazione parlata*, Tomo III, Napoli: Liguori, pp. 1696-714.

Voghera, M. and F. Cutugno (in press) AN.ANA.S: aligning text to temporal syntagmatic progression in Treebanks, in *Proceedings of the fifth Corpus Linguistics Conference*, Liverpool 20-23 July 2009.

I due volumi raccolgono i contributi presentati al Congresso Internazionale “La Comunicazione Parlata”, tenutosi a Napoli dal 23 al 25 febbraio 2009 ed organizzato dal G.S.C.P. (Gruppo di Studio sulla Comunicazione Parlata della Società di Linguistica Italiana). Il Convegno ha costituito un'occasione feconda di interazione e confronto tra studiosi di diversa provenienza e competenza, interessati alle dinamiche della comunicazione parlata nei suoi molteplici aspetti.

Il primo volume comprende i lavori relativi alla dimensione cognitiva (Parte I) e alla dimensione pragmatica e semantica (Parte II). Il secondo volume ospita i contributi relativi ai tradizionali settori dell'analisi linguistica, come la dimensione fonica (Parte III) e la dimensione morfosintattica (Parte IV), nonché quelli prodotti all'intersezione di più campi disciplinari, come antropologia e sociologia della comunicazione parlata (Parte V), comunicazione parlata e patologie del linguaggio (Parte VI), strumenti e tecniche per lo studio della comunicazione parlata (Parte VII).

Massimo Pettorino è professore di Comunicazione Parlata e di Linguistica generale presso l'Università degli Studi di Napoli L'Orientale ed è coordinatore nazionale del G.S.C.P.

Antonella Giannini è professore di Fonetica Sperimentale e di Linguistica generale presso l'Università degli Studi di Napoli L'Orientale e si occupa di prosodia e fonetica articolatoria e acustica.

Francesca M. Dovetto è professore di Glottologia e Linguistica presso l'Università degli Studi di Napoli Federico II. Si occupa di storia della linguistica, etimologia, lessico e morfologia dell’italiano e degli aspetti connotativi del parlato.