



SCUOLA
NORMALE
SUPERIORE
PISA

PER UNA CARATTERIZZAZIONE ACUSTICA DEL *CLEAR SPEECH*

(in corso di stampa negli Atti del Congresso Internazionale
La comunicazione parlata / Spoken Communication,
Napoli, 23-25.II.2006)
Silvia Calamai

Di certe frasi diceva il principio, di altre la fine;
borbottava delle parole incomprensibili,
altre le sillabava chiare.
(Italo Svevo, *Senilità*)

Abstract

La ricerca prende in esame alcuni aspetti acustici relativi al passaggio dal parlato 'normale' al parlato cosiddetto *clear* e affronta preliminarmente varie questioni di carattere terminologico e metodologico concernenti le indagini che si prefiggono di meglio comprendere questa modalità d'eloquio. I dati analizzati provengono da due parlanti di sesso maschile, uno italiano (Pisa), uno di lingua mòoré (Ougadougou, Burkina Faso). Per il locutore pisano sono stati osservati i valori relativi alla velocità d'eloquio e alla durata, in maniera da poter verificare se il parlato *clear* sia complessivamente più lento. Per entrambi i parlanti sono stati misurati i valori formantici delle vocali toniche e l'estensione dello spazio vocalico, in modo tale da appurare se esiste un legame tra l'estensione dell'inventario vocalico e la maggiore o minore tendenza verso il parlato *clear*.

1. Cos'è il parlato *clear*?

Prima di descrivere la modalità d'eloquio al centro dell'indagine ci pare opportuno affrontare alcuni aspetti di carattere terminologico: del determinante *clear* non crediamo sia possibile offrire una traduzione letterale, dal momento che *parlato chiaro* potrebbe essere equivocato con il parlato di chi si esprime "senza reticenze, in modo franco e fermo" (GDLI)¹. Non convince neppure l'alternativa 'parlato iperarticolato', dal valore parzialmente metonimico: il parlato *clear* è – per definizione – iperarticolato, ma

¹ Naturalmente, tra le accezioni di *chiaro* compare anche quella che si addice al nostro oggetto di studio – "che si ode distintamente, che ha un timbro limpido, puro (un suono, una voce)" – ma l'ambiguità di fondo rimane.

l'iperarticolazione è proprio una delle caratteristiche (o – se si preferisce – uno degli effetti) del parlato *clear*. In questa indagine, viene pertanto mantenuta la forma inglese.

Il contenitore che va sotto il nome di 'parlato *clear*' appare oltremodo vasto e racchiude al suo interno molte e differenti modalità di eloquio, tutte utilizzate in situazioni comunicative 'difficili': in contesti di rumore ambientale (il cosiddetto *Lombard speech*); in ambiente riverberante; nella comunicazione diretta al bambino, a soggetti non normoudenti, a stranieri; nel parlato ad alta voce; nel parlato emesso durante una comunicazione simultanea; nel parlato prodotto in condizioni di stress; nel parlato emesso a grandi distanze dall'ascoltatore². Questi vari stili di eloquio presentano molti fenomeni in comune e aree di elevata sovrapposizione: parlando a stranieri si adottano strategie usate con soggetti non normoudenti (come l'elevato volume della voce, un uso accentuato della gestualità, una minore velocità d'eloquio, etc.), oppure si attuano strategie documentate anche nel parlato ai bambini, come ad esempio la costruzione di frasi più brevi e meno complesse, l'introduzione di pause tra i costituenti, l'uso di termini generali e non specifici, le ripetizioni. Ma i fenomeni che ricorrono nel parlato *clear* non si ritrovano necessariamente in contemporanea in tutte le modalità d'eloquio *supra* succintamente delineate: alcuni appaiono più tipici e più frequenti di altri e ciascuna tipologia sarà caratterizzata da una diversa incidenza delle singole caratteristiche. In estrema sintesi, nel *clear speech* i parlanti tendono a produrre un eloquio che sia più facilmente riconoscibile e comprensibile: studi di tipo percettivo mostrano infatti un vantaggio percettivo di questa modalità rispetto ad altre tipologie, sia per ascoltatori normoudenti che per ascoltatori con deficit uditivi (Picheny, Durlach & Braida 1985; Uchanski *et alii* 1996). Si tratta dunque di un parlato più 'resistente' alle condizioni avverse: gli aggiustamenti che permettono una tale 'resistenza' hanno in parte un carattere universale e in parte appaiono motivati da restrizioni linguo-specifiche.

Gli studi sul *clear speech* cominciano ad apparire con una certa frequenza a partire dagli anni Ottanta, ma ciò che da un quarto di secolo è definito *clear speech* veniva prima indicato con le etichette di "buona dizione" e di "buona enunciazione" (Uchanski 2005: 208)³. Picheny, Braida, Lindblom, Moon e Bradlow hanno estesamente indagato questa tipologia d'eloquio, soprattutto in riferimento all'angloamericano; sono molto scarse le ricerche che si concentrano su altre lingue (disponiamo di dati relativi al francese, allo spagnolo e al croato), sono rarissimi gli studi interlinguistici⁴. La concentrazione di studi su un unico idioma rende peraltro arduo verificare

² Ma il parlato *clear* può anche non essere intenzionale: ci sono parlanti il cui eloquio è, in condizioni normali, (molto) più comprensibile rispetto a quello di altri. Il parlato *clear* prodotto in maniera intenzionale è peraltro acusticamente molto simile a quello prodotto inavvertitamente (vd. Bond & Moore 1994).

³ Queste divergenze nelle 'etichette' potrebbero celare a nostro avviso una difficoltà più profonda, su cui torneremo in § 2.

⁴ Un esempio è l'indagine di Smiljanić & Bradlow (2005), ove viene confrontato l'inglese con il croato.

con precisione l'esistenza di strategie universali attive in questa modalità d'eloquio. È comunque possibile – a partire dalla letteratura esistente – accennare alle principali modifiche attese nel passaggio da uno stile per così dire *normal* a uno stile *clear*. Da un punto di vista molto generale, rispetto a un eloquio 'normale', il parlato *clear* mostra una maggiore precisione articolatoria (e dunque una minore variabilità dovuta agli effetti della coarticolazione), una minore velocità d'eloquio, un maggior numero di pause, un aumento del volume. Si tratta pertanto di un parlato più intenso (l'aumento ammonta a circa 5-8 dB); l'innalzamento del volume ha come effetto secondario l'aumento della prima formante per le vocali⁵, i *glides*, le liquide e le nasali. La maggiore lentezza d'eloquio è dovuta sia a un più alto numero di pause, sia a un maggior numero di inserzioni di foni. La differenza nella velocità d'eloquio tra stile cosiddetto 'conversazionale' e stile *clear* è un dato che emerge in molti studi e la maggiore lentezza rappresenta uno dei caratteri maggiormente distintivi del parlato *clear* (Uchanski *et alii* 1996: 495). Tuttavia – è bene ricordarlo – indagini più recenti hanno mostrato come il parlato *clear* abbia caratteristiche acustiche indipendenti dalla velocità d'eloquio: con informatori particolarmente addestrati è infatti possibile produrre parlato *clear* anche a velocità di eloquio più elevate (Krause & Braida 2002)⁶. La maggiore precisione articolatoria ha come effetti acustici una maggiore durata del V.O.T., una maggiore estensione dello spazio vocalico, una maggiore durata delle transizioni formantiche. Per gli aspetti soprasegmentali, si registra un valore più elevato della frequenza fondamentale e una maggiore modulazione. Il profilo del parlante ideale in grado di produrre parlato con un altissimo livello di intelligibilità è così tratteggiato da Bradlow, Torretta & Pisoni (1996: 270):

“a female who produces sentences with a relatively wide range in fundamental frequency, employs a relatively expanded vowel space that covers a broad range in F1, precisely articulates her point vowels, and has a high precision of intersegmental timing”.

Sembra dunque che il parlato *clear* sia caratterizzato da strategie che rafforzano sia il segnale, che diventa acusticamente più saliente, sia il codice, attraverso l'amplificazione della distanza acustica tra categorie. Tuttavia, come sottolinea Uchanski (2005), non sappiamo quanto e come questi fenomeni contribuiscono alla maggiore intelligibilità del parlato *clear* poiché non esistono esperimenti in cui ciascuno di essi sia manipolato in maniera sistematica. Non dobbiamo scordare infatti come le caratteristiche del *clear speech* siano sempre definite in relazione a un'altra modalità d'eloquio (forme di citazione, parlato semispontaneo, o spontaneo, o casuale, o

⁵ Ma disponiamo di risultati sperimentali che vanno anche in senso contrario (vd. § 3.).

⁶ Questo risultato non desta particolari sorprese, se si pensa al cammino di 'revisione' che ha subito il modello del cosiddetto *target-undershoot* di Lindblom: i parlanti possono raggiungere i loro bersagli acustici anche a velocità d'eloquio sostenute, poiché la velocità d'eloquio non risulta essere un buon predittore per il raggiungimento (o il mancato raggiungimento) dei bersagli.

conversazionale...), quindi abbiano un carattere *relativo* e non assoluto⁷. Le osservazioni avanzate in § 2.1. intendono pertanto sottolineare alcuni aspetti di questa sorta di ‘cortocircuito metodologico’.

2. L’indagine sperimentale

2.1. *Appunti sul metodo*

Abbiamo già sottolineato in § 1. come dietro l’iperonimo di parlato *clear* si celi in realtà una multiforme varietà di parlato, e rileviamo ora come sia un problema non secondario ottenere parlato *clear* in condizioni sperimentalmente controllate.

In letteratura sono usati sia metodi espliciti che metodi impliciti⁸. Le metodologie dirette prevedono la creazione di una situazione ambientale che permetta di ottenere questa modalità con strategie che potremmo definire ‘ecologiche’: è il contesto ambientale – sapientemente costruito dal ricercatore – a innescare la modalità *clear*. Ad esempio, in Clark *et alii* (1987) l’eloquio ‘normale’ viene ottenuto chiedendo ai soggetti di leggere le frasi “with normal, or unmarked, articulatory effort” (163)⁹, mentre il parlato *clear* viene elicitato facendo indossare ai soggetti cuffie che trasmettono rumore e chiedendo loro di comunicare con un ascoltatore fuori dalla camera silente; per l’analisi delle caratteristiche acustiche del parlato prodotto in presenza di rumore, Junqua (1993) confronta il parlato registrato in assenza di rumore e il parlato prodotto dai medesimi parlanti mentre questi ascoltano rumore bianco gaussiano di 85 dB SPL attraverso cuffie calibrate; le parole sono elicitate in isolamento e possiamo supporre che siano lette.

Nelle metodologie di tipo indiretto, si chiede esplicitamente al soggetto di “iperarticolare” (le istruzioni più frequenti sono riportate *infra*), ovvero si chiede di ‘far finta’ di parlare a un ascoltatore con una ridotta capacità di comprensione (perché non normoudente, perché parlante non nativo): il *clear*

⁷ La difficoltà insita nel confrontare una modalità d’eloquio ‘marcata’ rispetto a una genericamente ‘non marcata’ (ma non sappiamo in che misura) è stata sottolineata da Ferguson (1981) proprio nell’ambito dello studio sul parlato a stranieri: nella sua rassegna oppone il *Foreigner Talk* al cosiddetto *baseline register* (“the ordinary, conversational language of the community”, altrimenti detto “normal talk”, “adult speech”, “native talk”).

⁸ Nella maggior parte degli studi che analizzano il parlato *clear* il fattore ‘scatenante’ è sempre di tipo ambientale, mentre un altro filone di indagine si concentra sulle caratteristiche di tipo inerentemente linguistico in grado di influenzare le strategie di produzione: si hanno produzioni più accurate anche quando il parlante introduce informazione nuove nel discorso (vd. tra gli altri Bolinger 1963; Hawkins & Warren 1994), o – nei casi di competizione lessicale – quando il parlante pronuncia parole meno intelligibili rispetto a parole più intelligibili (*easy* vs. *hard words*, nella terminologia di Wright 2003).

⁹ Ci pare degna di nota la seguente riflessione a margine: “This is the classic ‘open loop’ data gathering experiment of the kind commonly done in phonetics. It was open loop in the sense that the speaker had no communication goal in the form of a listener with whom to interact”.

speech appare dunque assimilabile, per certi versi, al parlato recitato¹⁰, dal momento che ai soggetti viene richiesta una sorta di ‘messa in scena’. In Picheny *et alii* (1985: 97) la strategia adottata è la seguente:

“When recording conversational speech, the talker was instructed to *recite*¹¹ the materials in the same manner in which he spoke in ordinary conversation. [...] When recording clear speech, the talker was instructed to speak as clearly as possible, as if he were trying to communicate in a noisy environment or with an impaired listener¹². The talkers were also told to enunciate consonants more carefully and with greater (vocal) effort than in conversational speech and to avoid slurring words together.”

In Lindblom & Moon (1988: 23) i soggetti devono leggere parole isolate “at comfortable vocal efforts and speaking tempos” così da ottenere forme di citazione, e leggere le medesime parole iperarticolarlo, come se si rivolgersero a un ascoltatore non nativo. Ai soggetti di Moon & Lindblom (1994: 43) viene chiesto di ‘iperarticolare’, ovvero di leggere le parole nella maniera più chiara possibile; per evitare che questa modalità venga per così dire ‘disattivata’, la lettura viene talvolta interrotta dal ricercatore il quale finge che la parola appena pronunciata non sia stata compresa e chiede pertanto ai soggetti di ripeterla: in questo modo, la procedura permette ai soggetti di mantenere la modalità *clear* (e consente di ottenere anche stimoli classificati come “extra clear”). Nell’indagine di Bradlow (2002: 247):

“For the conversational speaking mode, subjects were told to read at their normal pace without any particular attention to clarity. Subjects were told to *imagine*¹³ that the intended listener of these recordings was someone highly familiar with her or his voice and speech patterns. For the clear speaking mode, subjects were told to read the sentences as if speaking to a listener with a hearing loss or from a different language background.”

In Ferguson & Kewley-Port (2002), il soggetto unico legge una lista “using his normal, conversational speaking style” e in un altro momento legge lo stesso materiale come se parlasse a un soggetto non normoudente, avendo cura di pronunciare ciascuna parola con attenzione. In Perkell *et alii* (2002: 1629), il parlato ‘normale’ viene ottenuto “by asking the subjects to pronounce the utterances at a ‘conversational’ pace”, mentre il parlato *clear* viene registrato “by telling the subjects that someone in the next room was checking their pronunciation and they would be rewarded according to the number of

¹⁰ In molte indagini sperimentali, il parlato analizzato è davvero una forma (particolare) di parlato recitato: per quanto riguarda il *foreigner talk* si vedano ad esempio le istruzioni di Ferguson (1975), poi riprese nelle indagini di Gaetano Berruto e di Stephan Schmid sull’italiano a stranieri, in cui si chiede ai soggetti di modificare alcune frasi come se stessero rivolgendosi a un parlante non nativo con scarse conoscenze della lingua oggetto d’analisi (vd. almeno Schmid 2003).

¹¹ Corsivo nostro.

¹² Si assume, implicitamente, che il parlare in un ambiente rumoroso metta in atto strategie simili a quelle attivabili in una comunicazione con soggetti non normoudenti.

¹³ Corsivo nostro.

utterances pronounced correctly”. In Krause & Braidà (2002: 2166) si chiede esplicitamente l’imitazione dell’imitazione:

“In order to obtain a form of clear speech from each talker with minimal training, the talkers were familiarized with the characteristics of previously obtained clear speech. The talkers listened to samples of both conversational and clear speech materials recorded from an earlier study of clear speech (Picheny *et alii*, 1985), and differences between the two speaking modes were discussed. The talkers were asked to *mimic*¹⁴ the clear speech that had been presented, and they were given feedback by the experimenter on both rate and clarity.”

Come mostra questo florilegio di citazioni, le indicazioni più frequenti assomigliano a quelle pronunciate da un regista a un attore (“fai finta di parlare a un sordo”, “fai finta di parlare a uno straniero”, “fai finta di parlare in condizioni di rumore ambientale”, “cerca di pronunciare con attenzione e con chiarezza ogni parola”...) e in qualche modo celano un’assunzione implicita (e certamente da verificare) relativa al fatto che si possa ottenere una *medesima* modalità d’eloquio a partire da istruzioni così diverse. Tuttavia persino una medesima istruzione impartita a soggetti differenti può produrre esiti diversi, come si rileva in Perkell *et alii* (2002: 1640): “There could be differences in the way the subjects interpreted the instructions and they could differ in the way they produce and/or perceive the acoustic correlates of clear speech”.

Anche un’annotazione a margine dell’esperimento di Smiljanić & Bradlow (2005: 1688) ci pare celi per certi versi un limite nel disegno sperimentale – limite peraltro difficile da arginare:

“It is important to note that both ‘conversational’ and ‘clear’ speech terms refer to read laboratory speech elicited by specific instruction given to talkers rather than in a more naturalistic situation. Although, ultimately, we would like to investigate spontaneously produced clear and conversational speaking styles, we believe that data obtained in this study approximate this goal sufficiently since talkers implemented conversational-to-clear speech articulatory modifications which resulted in significant intelligibility benefits.”

Tutte le riflessioni *supra* trascritte si riferiscono nella grande maggioranza dei casi a parlato letto. Assumiamo – non senza qualche difficoltà – che esso rappresenti uno stile di *default* e che sia possibile, al suo interno, confrontare uno stile per così dire ‘normale’ (*normal speech* è l’etichetta più frequente negli studi¹⁵) con uno stile, appunto, *clear*. È legittimo a questo punto chiederci se

¹⁴ Corsivo nostro.

¹⁵ Talvolta l’opposizione è tra forme *clear* e forme di citazione (vd. ad esempio Moon & Lindblom 1994). Ma l’etichetta ‘forma di citazione’ non è esente da contraddizioni, come mostra la riflessione di Koreman (2006: 582), che suggerirebbe indirettamente quasi un’identità (o, almeno, una prossimità molto pericolosa) tra forme di citazione e forme *clear*: “To elicit a word in its canonical form, speakers may be requested to produce the word in citation form, i.e., the segment sequence which is produced when the word is articulated clearly and in isolation. But at the same time it is not certain that the citation form is really a good indicator of the canonical form: it may be overarticulated in the sense of a “spelling pronunciation””. Per

normal sia davvero uno stile ‘normale’ o se non sia piuttosto una forma di citazione, già di per sé in qualche modo iperarticolata, e dunque, in qualche modo, sempre *clear*. Se guardiamo al passaggio da *normal* a *clear* come a un *continuum* entro il quale è possibile individuare facilmente (?) gli estremi, non possiamo non scorgere qualche aporia nel creare opposizioni ‘da laboratorio’ che, pur all’interno di un quadro enormemente semplificato rispetto alla vita reale, presentano comunque aspetti metodologicamente ambigui. Un’ambiguità di non poco conto risiede a nostro avviso nell’ottenere alla fine un parlato recitato assumendo che questo sia *clear* (in molti disegni sperimentali si allude proprio a un lessico della scena: *imagine, recite, rehearse, mimic...*)¹⁶.

2.2. I materiali

Questa indagine pilota presenta osservazioni soprattutto intra-parlante e intende offrire alcuni dati relativi a due locutori, uno di lingua italiana (proveniente da Pisa) e uno di lingua móorè (proveniente da Ouagadougou), per quanto concerne alcuni aspetti acustici del parlato *clear*. In entrambi i casi le analisi qui riportate fanno parte di ricerche più ampie, in corso d’opera¹⁷.

Le due lingue prescelte sono tipologicamente molto differenti¹⁸: l’obiettivo di lungo termine è quello di valutare se esistano strategie comuni nel passaggio dall’uno all’altro stile, in lingue che differiscono anche in misura sostanziale per quanto concerne l’inventario fonologico.

Entrambi i locutori hanno prodotto frasi con struttura “(non hai capito) ho detto X (non ho detto y)” e hanno poi ripetuto la parola bersaglio X in isolamento¹⁹. Il soggetto pisano ha inoltre letto 30 frasi sintatticamente corrette, ma semanticamente impossibili²⁰. A entrambi i parlanti è stato detto di “far finta di parlare a una persona straniera”, e l’unico suggerimento supplementare eventualmente fornito – di fronte a dubbi, esitazioni, perplessità – è stato quello di “parlare in modo da essere capiti il più facilmente possibile”. Come in molti studi incentrati su questa modalità d’eloquio, ci siamo limitati, per il momento, a un solo parlante per lingua: è peraltro ben documentata in letteratura l’esistenza di strategie altamente

non parlare dell’etichetta *conversational*, generalmente usata in riferimento a un tipo di parlato che quasi sempre è parlato letto, come mostrano alcune delle citazioni *supra* riportate.

¹⁶ Ma, a ben guardare, non sono ancora ben note le caratteristiche acustiche del parlato recitato e un percorso di ricerca in tal senso appare molto suggestivo (cfr. Boulakia 2002: 75).

¹⁷ Precisiamo che è stata compiuta solo in minima parte l’analisi statistica: l’indagine ha un carattere esplorativo e i confronti sono di carattere qualitativo.

¹⁸ Il móoré appartiene alla ramificazione settentrionale del gruppo Gur (voltaico), compreso entro la famiglia Niger-Congo.

¹⁹ La medesima metodologia è usata in Clark *et alii* (1987: 163) e in Lindblom & Moon (1988: 23).

²⁰ Queste frasi si rendono dunque disponibili per un test percettivo: l’assenza di un contesto semanticamente plausibile non fornisce aiuti supplementari all’ascoltatore nel compito di identificazione.

individuali nella produzione di parlato *clear*²¹. Del resto, una variabile non secondaria ci pare, a nostro avviso, quella rappresentata dall'effetto 'addestramento': in alcune registrazioni pilota, parlanti 'non addestrati' hanno prodotto parlato semplicemente 'urlato' (e i materiali sono stati ovviamente scartati, poiché inutilizzabili). Non sarà probabilmente un caso che la scelta si sia poi orientata su parlanti non ingenui (un attore di teatro vernacolo per l'italiano parlato a Pisa, un linguista per il mòoré). Anche il soggetto di Ferguson & Kewley-Port (2002: 260) non è un soggetto che definiremmo 'ordinario': "the talker, a 59-year-old audiologist, was chosen for his dialect [...] and for his extensive experience communicating with hearing-impaired people". Sulla stessa linea anche l'indagine di Krause & Braida (2002), che peraltro dedica molto spazio alle metodologie con cui sono ottenuti i dati:

"In order to improve the chances of obtaining clear speech at normal speaking rates [...] naturally, much attention was given to talker selection and training. [...] only talkers with significant public speaking experience (e.g. students or professionals in television or radio broadcasting, public speaking, or other communications disciplines) were considered" (2166).

La 'selezione' del parlante e il suo 'addestramento' appaiono dunque elementi di rilievo per l'analisi. Sembra che parlanti 'normali' non siano buoni soggetti per l'analisi di parlato *clear*.

3. Alcuni tratti del parlato *clear*

Per quanto concerne l'italiano parlato a Pisa, in relazione alle frasi semanticamente impossibili, sono state compiute le seguenti verifiche strumentali:

- (a) presenza/assenza di pause;
- (b) maggiore/minore lentezza d'eloquio.

Per quanto riguarda (a), le frasi del parlato *clear* non sono state prodotte con l'inserimento di pause²². Per quanto concerne (b), le frasi lette in modalità *clear* sono complessivamente più lente, e la differenza è statisticamente significativa (Tabella 3.1.)²³.

²¹ "Different talkers vary in their overall intelligibility and their ability to speak clearly" (Picheny *et alii* 1985: 101); "[subjects] could differ in the way they produce and/or perceive the acoustic correlated of clear speech" (Perkell *et alii* 2002: 1640); "only one talker was examined, and talkers vary considerably in their ability to produce highly intelligible speech" (Krause & Braida 2002: 2166); "talkers may employ different strategies for producing clear speech at normal rates" (Krause & Braida 2004: 376); "there is significant individual variability in the characteristics of "intrinsically clear" speech across talker" (Hazan & Markham 2004: 3117).

²² Ad eccezione di un solo caso, con una pausa di 61 ms tra il SV e il SN. Questo risultato potrebbe essere dovuto alla relativa brevità delle singole frasi, prive di subordinate. Sarà dunque necessario ripetere l'esperimento con batterie costituite da brevi brani, sintatticamente più complessi.

²³ Anova a una via: velocità d'eloquio, sig. .000 (F = 54.688); durata media, sig. .000 (F = 59,715). Per la modalità *normal* i valori minimo e massimo della velocità d'eloquio sono,

Tabella 3.1. - *Normal vs clear*: valori medi e deviazione standard della velocità d’eloquio e della durata.

Stile	Velocità d’eloquio (sill/sec)	Durata media
<i>normal</i>	5.52 (.76)	2.06 (.29)
<i>clear</i>	4.19 (.56)	2.71 (.33)

L’assenza di pause nella modalità *clear* peraltro fa coincidere la velocità di articolazione (*articulation rate*) con la velocità di elocuzione (*speech rate*), in cui sarebbero comprese anche le pause²⁴. In assenza di pause, una ulteriore indagine dovrà riguardare l’entità dell’allungamento per ciascun fono, ai fini di verificare quali classi siano maggiormente ‘allungabili’²⁵. La Tabella 3.2. mostra a titolo meramente esemplificativo i risultati delle misurazioni di durata (in ms) relative alla frase *Le fitte mangiano sotto di lui*:

Tabella 3.2. - *Normal vs clear*: confronto fra valori di durata

Stile	[l]	[e]	[f]	[i]	[t]	[e]	[m]	[a]	[n]	[dʒ]	[a]	[n]	[o]	[s]	[o]	[t]	[o]	[d]	[i]	[l]	[ui]
<i>Norm</i>	34	69	143	89	148	56	66	125	50	46	73	25	64	79	60	115	50	39	71	38	279
<i>Clear</i>	76	105	158	118	294	142	128	158	66	60	64	47	54	130	78	164	58	62	69	67	283

Audio normale 

Audio clear 

Ma visto che il parlato *clear* non è semplicemente un parlato prodotto con uno stile d’eloquio più lento, abbiamo osservato anche altri parametri acustici, riferiti al sistema vocalico. Le vocali sono i suoni più udibili alle grandi distanze e in contesti di rumore ambientale: parrebbe pertanto che le vocali toniche all’interno di una parola forniscano il contributo maggiore alla intelligibilità della parola stessa²⁶.

Due aspetti dei vocalismi italiano e mòoré sono stati indagati: i valori formantici e l’estensione dello spazio vocalico. Per quanto concerne la struttura formantica, intendiamo verificare dove si dirigono le formanti nel passaggio da *normal* a *clear*. A questo proposito è stato confrontato il valore della vocale

rispettivamente, 4.058 e 7.027; per la modalità *clear* sono, rispettivamente, 3.215 e 5.173. Per quanto concerne la durata complessiva delle frasi, nella modalità *normal* i valori minimo e massimo sono, rispettivamente, 1.423 ms e 2.878 ms; nella modalità *clear* sono, rispettivamente, 1.933 ms e 3.415 ms.

²⁴ Un’agile rassegna concernente le modalità di misurazione del tempo della voce è Roach (1998).

²⁵ Replicando, *mutatis mutandis*, l’esperimento di Fónagy & Fónagy (1966).

²⁶ Così Fónagy & Fónagy (1966), Schulman (1989), Aylett (1999). A differenti conclusioni porta Cutler (2002).

tonica dell'*item* in isolamento con il valore della vocale tonica della parola bersaglio X nella struttura “(non hai capito) ho detto X (non ho detto y)”²⁷.

Nei grafici seguenti riportiamo i valori medi di tre ripetizioni per ciascuna categoria vocalica. La freccia indica la direzione dello spostamento, da *normal* a *clear*.

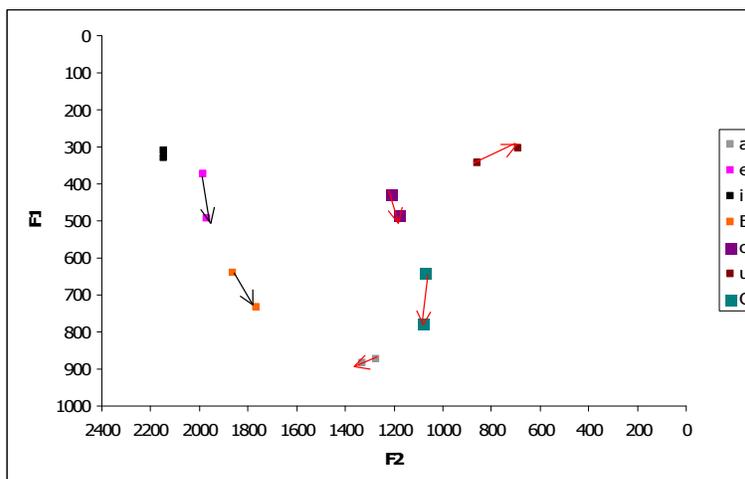


Figura 3.1. - Vocalismo tonico, parlante italiano (Pisa)

Per /a/ si ha un innalzamento della prima e della seconda formante. Le vocali estreme tendono a presentare una diminuzione della prima formante; per la seconda formante si registra una maggiore stabilità per /i/ e un consistente abbassamento per /u/²⁸. Per entrambe le vocali medio-basse, si ha un consistente aumento della prima formante e una diminuzione della seconda (rilevante per la vocale anteriore, irrilevante per quella posteriore). Anche per le vocali medie l'andamento è simile: un aumento per la prima formante e una diminuzione per la seconda.

²⁷ Questa frase è stata tradotta in mòoré con l'aiuto di Norbert Nikiéma, che ringrazio.

²⁸ La maggiore stabilità di /i/ rispetto alle vocali posteriori è ampiamente nota in letteratura.

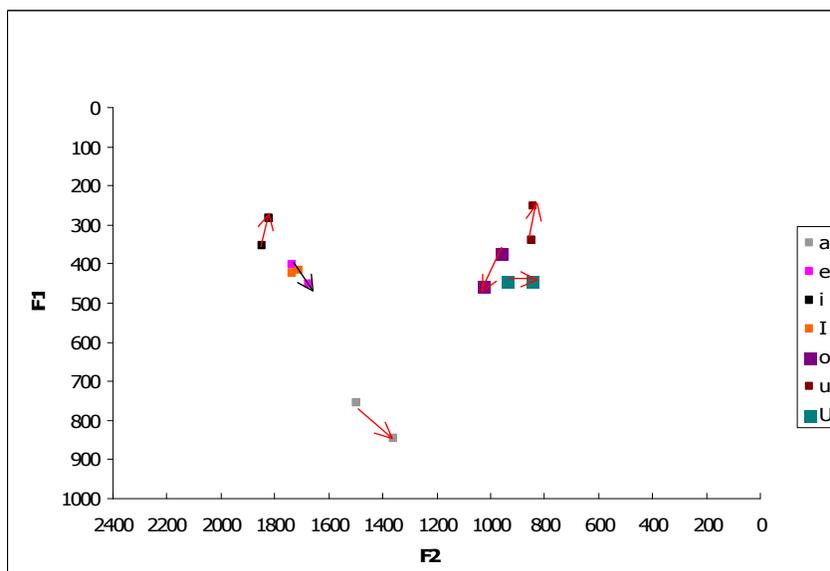


Figura 3.2. - Vocalismo tonico breve, parlante m̀oré (Ougadougou)

Il sistema vocalico della lingua m̀oré è molto più complesso rispetto a quello italiano: ciascun fonema vocalico può essere lungo o breve, orale o nasale, e può essere di tono alto, basso o medio. I dati acustici qui presentati si riferiscono al sistema vocalico orale, lungo e breve, composto di sette timbri vocalici, per un totale di quattordici fonemi: una vocale centrale /a/, due vocali medie /e o/, quattro vocali alte che – sia per la serie breve sia per la serie lunga – presentano una opposizione nel tratto [\pm ATR]: /i ɪ u U/²⁹.

In buona parte, i comportamenti registrati per il sistema vocalico italiano trovano conferma anche nel caso di una lingua dall'inventario più affollato. Nella presentazione dei risultati, le medie delle vocali brevi (Figura 3.2.) precedono quelle delle vocali lunghe (Figura 3.3.).

²⁹ In realtà, l'inquadramento fonologico del vocalismo m̀oré presenta alcuni aspetti particolarmente controversi, che non è il caso di affrontare in questa sede (rimandiamo pertanto a Calamai & Bertinetto 2006). Nei due grafici, le vocali lunghe sono indicate mediante la ripetizione della vocale; i simboli /ɪ ɪ u U/ si riferiscono alle vocali [-ATR].

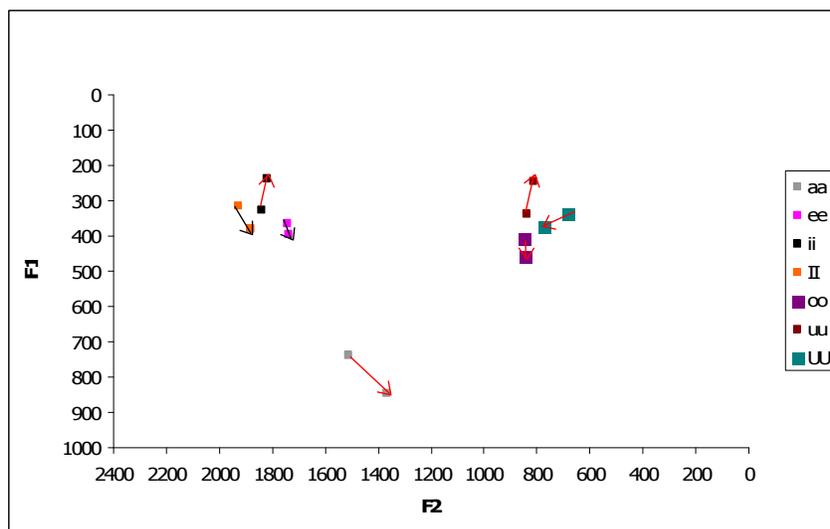


Figure 3.3. - Vocalismo tonico lungo, parlante mòoré (Ougadougou)

Per la vocale /a/, sia breve che lunga, si registra una diminuzione della prima formante e un aumento della seconda. Come per l'italiano, le vocali alte presentano un abbassamento della prima formante, sia sull'asse anteriore che sull'asse posteriore. Per le vocali alte [-ATR] il movimento è consistente solo nella serie delle vocali lunghe ma risulta per certi versi contraddittorio³⁰. Un comportamento discordante è offerto dalle vocali medie: la vocale anteriore tende a presentare un innalzamento della prima formante e un abbassamento della seconda; quella posteriore ha una diminuzione delle due formanti solo per la vocale breve.

Dunque non per tutte le vocali si ha un aumento della prima formante, a differenza di quanto viene rilevato in Ferguson & Kewley-Port (2002: 262-263). Il dato non stupisce: la prima formante aumenta sì con l'aumentare dello sforzo vocale (Lienard & Di Benedetto 1999), ma per le vocali estreme la vocale iperarticolata presenta tendenzialmente una F1 più bassa³¹.

Per valutare se lo spazio vocalico è sempre più ampio nel parlato *clear* è stata calcolata l'area del poligono formato dalle linee che congiungono le

³⁰ Per altri versi, una mancanza di risultati univoci è stata rilevata anche nel confronto tra vocali tese e vocali rilassate in studi relativi alla lingua inglese (vedi Picheny *et alii* 1986: 441): sembra che le vocali rilassate siano più sensibili ai cambiamenti nella modalità d'eloquio rispetto a quelle tese, ma il legame tra frequenze formantiche e stile è definito senza mezzi termini "ambiguous".

³¹ Per il parlato ad alta voce, invece, non è esattamente così: i dati di Schulman (1989), Junqua (1993, 1996), Castellanos *et alii* (1996) mostrano un aumento più o meno consistente per tutte le formanti. Una prova ulteriore – crediamo noi – del fatto che il parlato *clear* non sia semplicemente un parlato prodotto a un volume di voce più alto. Nel caso delle vocali posteriori l'iperarticolazione tende a enfatizzare l'arrotondamento labiale, che tra le sue conseguenze acustiche ha proprio l'abbassamento di F1.

vocali rappresentate su un piano con le coordinate F1/F2 (Pätzold & Simpson 1997). Il dato è confermato per entrambe le lingue, che pure presentano un inventario vocalico molto diverso tra loro³² (Tabella 3.3.): la distanza acustica tra singole vocali tende ad aumentare nel parlato *clear*.

Tabella 3.3. - *Normal vs clear*: aree vocaliche a confronto

Parlante	<i>normal</i>	<i>clear</i>
Italiano (Pisa)	344.417 Hz ²	439.096 Hz ²
Mòoré (V brevi)	192.483 Hz ²	294.258 Hz ²
Mòoré (V lunghe)	218.053 Hz ²	388.580 Hz ²

I valori medi riportati sul piano cartesiano nelle Figure 3.4., 3.5. e 3.6. mostrano l'estensione dei cambiamenti spettrali nel passaggio da *normal* a *clear*:

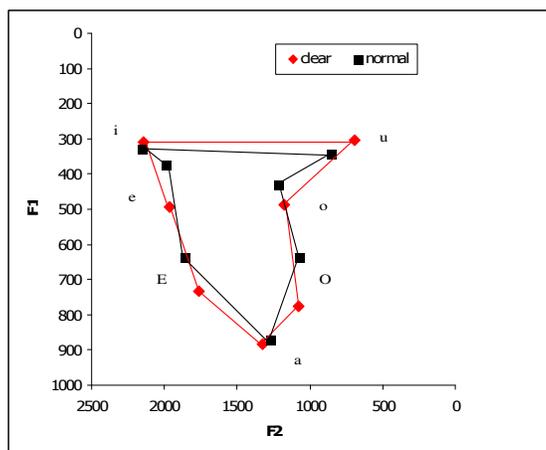


Figura 3.4. - Vocalismo italiano

³² Risultati simili sono in Bradlow (2002) nel confronto tra spagnolo e inglese, e in Smiljanić & Bradlow (2005) nel confronto tra inglese e croato. In Hazan & Markham (2004) sono presentati e discussi dati di parlato *clear* che, al contrario, *non* presentano una maggiore dispersione vocalica.

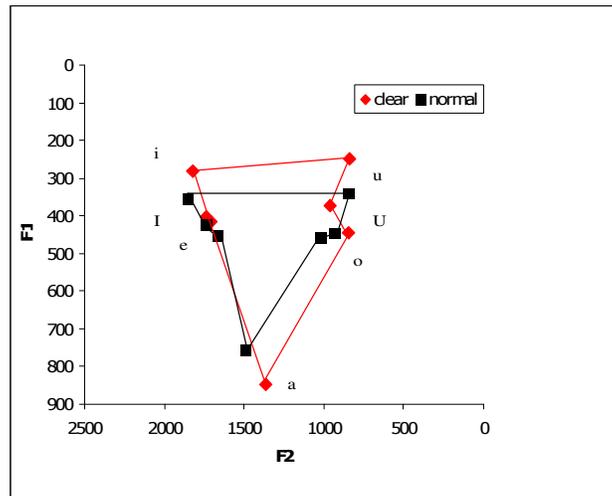


Figura 3.5. - Vocalismo mòré (V brevi)

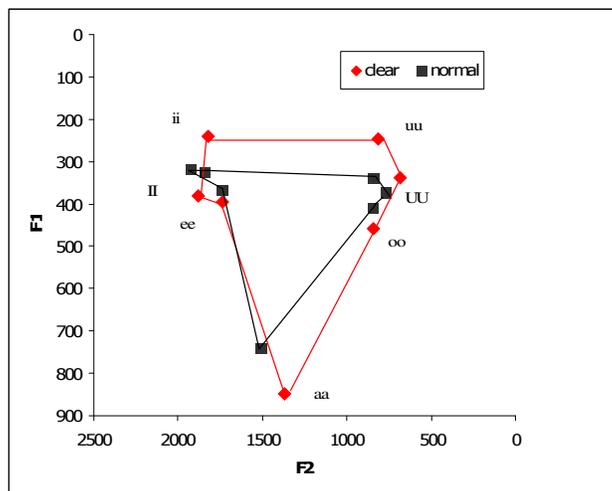


Figura 3.6. - Vocalismo mòre (V lunghe)

A questa maggiore estensione dello spazio vocalico, indipendentemente dal numero dei fonemi coinvolti, Bradlow (2002: 261-262) ha dato la seguente spiegazione:

“the focus of the acoustic-phonetic enhancements in clear speech production is not determined by the system of underlying phoneme contrasts of the language. Rather, clear speech is characterized by global hyperarticulation”.

Per le vocali /i a u/, si registra una simile espansione anche nel *baby talk* di madri inglesi, russe e svedesi (Kuhl *et alii* 1997) e nel parlato a stranieri di

soggetti inglesi (Uther *et alii* 2007): in entrambi i casi si tratterebbe di un espediente didattico che consente di ‘esagerare’ i contrasti fonetici rilevanti enfatizzando i prototipi acustici estremi.

Senza dubbio, come è stato in più luoghi suggerito, si rendono necessarie analisi ulteriori, e soprattutto nuovi confronti interlinguistici, per fornire prove supplementari al fatto che l’espansione dello spazio vocalico sia nel parlato *clear* un fenomeno indipendente dall’inventario fonologico. Inoltre, il peso dei fattori idiosincratici ribadito con forza da molti ricercatori ci spingerebbe non ad ampliare il numero dei parlanti ma a osservare con più profondità e con maggiori dettagli il singolo soggetto, per valutare se la variabilità del parlato *clear* sia in qualche modo legata al modo in cui viene prodotto il parlato ‘normale’, se – in altri termini – le caratteristiche del parlato *clear* possano essere in qualche modo predette da quelle che individuano, all’interno di ciascun parlante, la modalità ‘normale’ (cfr. Junqua 1996). Ma, crediamo noi, sarà necessaria anche una riflessione sul metalinguaggio adottato nei protocolli sperimentali, per meglio definire i limiti che potrebbero celarsi dietro risultati raggiunti nello studio di una modalità d’eloquio forse ancora particolarmente difficile da trattare.

Ringraziamenti

Si desidera ringraziare Federico Albano Leoni, Olle Engstrand, Antonella Giannini, Barbara Gili Fivela, Irene Ricci, Stephan Schmid.

Bibliografia

- Aylett M., 1999, *Modelling clarity change in spontaneous speech*, in Baddeley R.J., Hancock P.J.B, Földiák P. (a cura di), *Information theory and the brain*, New York, Cambridge University Press, 204-220.
- Bolinger D., 1963, *Length, vowel, juncture*, «Linguistics», 1, 5-29.
- Bond Z.S., Moore T.J., 1994, *A note on the acoustic-phonetic characteristics of inadvertently clear speech*, «Speech Communication», 14, 325-337.
- Boulakia G., 2002, *Linguistica e fonetica: senza voce o mezza voce?*, in De Dominicis A. (a cura di), *La voce come bene culturale*, Roma, Carocci, 63-85.
- Bradlow A.R., 2002, *Confluent talker- and listener-related forces in clear speech production*, in Gussenhoven C., Warner N. (a cura di), *Laboratory Phonology*, Berlin, Mouton de Gruyter, 7, 241-73.
- Bradlow A.R., Torretta G.M., Pisoni D.B., 1996, *Intelligibility of normal speech I: Global and fine-grained acoustic-phonetic talker characteristics*, «Speech Communication», 20, 255-272.
- Calamai S., Bertinetto P.M., 2006, *Sul vocalismo orale della lingua mòoré*, in Savy R., Crocco C. (a cura di), *Atti del II Convegno Nazionale AISV Associazione Italiana di Scienze della Voce Analisi prosodica. Teorie*,

- modelli e sistemi di annotazione*, Salerno, 30.XI-2.XII.2005, Torriana, EDK Editore, 419-438.
- Castellanos A., Benedí J.-M., Casacuberta F., 1996, *An analysis of general acoustic-phonetic features for Spanish speech produced with the Lombard effect*, «Speech Communication», 20, 23-35.
- Clark J.E., Lubker J.F., Hunnicut S., 1987, *Some preliminary evidence for phonetic adjustment strategies in communication difficulty*, in Steele R., Threadgold T. (a cura di), *Language Topics. Essays in honour of Michael Halliday*, Amsterdam/Philadelphia, Benjamins, II, 161-180.
- Cutler A., 2002, *Phonological processing: Comments on Pierrehumbert, Moates et alii, Kubozono, Peperkamp & Dupoux, and Bradlow*, in Gussenhoven C., Warner N. (a cura di), *Laboratory Phonology*, Berlin, Mouton de Gruyter, 7, 275-296.
- Ferguson C.A., 1975, *Towards a characterization of English foreigner talk*, «Anthropological Linguistics», 17, 1-14.
- Ferguson C.A., 1981, *'Foreigner Talk' as the name of a simplified register*, in Clyne M.G., Fishman J.A. (a cura di), *Foreigner Talk. International Journal of the Sociology of Language*, 28, 9-18.
- Ferguson S.H., Kewley-Port D., 2002, *Vowel intelligibility in clear and conversational speech for normal-hearing and hearing-impaired listeners*, «The Journal of the Acoustical Society of America», 112, 259-271.
- Fónagy I., Fónagy J., 1966, *Sound pressure level and duration*, «Phonetica», 15, 14-21.
- Hazan V., Markham D., 2004, *Acoustic-phonetic correlates of talker intelligibility for adults and children*, «The Journal of the Acoustical Society of America», 116, 3108-3118.
- Hawkins S., Warren P., 1994, *Phonetic influences on the intelligibility of conversational speech*, «Journal of Phonetics», 22, 493-511.
- Junqua J.-C., 1993, *The Lombard reflex and its role on human listeners and automatic speech recognizers*, «The Journal of the Acoustical Society of America», 93, 510-524.
- Junqua J.-C., 1996, *The influence of acoustics on speech production: A noise-induced stress phenomenon known as the Lombard effect*, «Speech Communication», 20, 13-22.
- Koreman J., 2006, *Perceived speech rate: The effects of articulation rate and speaking style in spontaneous speech*, «The Journal of the Acoustical Society of America», 119, 582-596.
- Krause J.C., Braida L.D., 2002, *Investigating alternative forms of clear speech: The effects of speaking rate and speaking mode on intelligibility*, «The Journal of the Acoustical Society of America», 112, 2165-72.
- Krause J.C., Braida L.D., 2004, *Acoustic properties of naturally produced clear speech at normal speaking rates*, «The Journal of the Acoustical Society of America», 115, 362-78.

- Kuhl P., Andruski J., Chistovich L., Kozhevnikova E., Ryskina V., Stolyarova E., Sundberg U., Lacerda F., 1997, *Cross-language analysis of phonetic units in language addressed to infants*, «Science», 277, 684-686.
- Lienard J.S., Di Benedetto M.G., 1999, *Effect of vocal effort on spectral properties of vowels*, «The Journal of the Acoustical Society of America», 106, 411-422.
- Lindblom B., Moon S.-J., 1988, *Formant Undershoot in Clear and Citation-form Speech*, «PERILUS», 8, 21-33.
- Moon S.-J., Lindblom B., 1994, *Interaction between Duration, Context, and Speaking Style in English Stressed Vowels*, «The Journal of the Acoustical Society of America», 96, 40-55.
- Pätzold M., Simpson A.P., 1997, *Acoustic Analysis of German Vowels in the Kiel Corpus of Read Speech*, «Arbeitsberichte des Instituts für Phonetik und digitale Sprachverarbeitung Universität Kiel», 32, 215-247.
- Payton K.L., Uchanski R.M., Braida L.D., 1994, *Intelligibility of conversational and clear speech in noise and reverberation for listeners with normal and impaired hearing*, «The Journal of the Acoustical Society of America», 95, 1581-1592.
- Perkell J.S., Zandipour M., Matthies M.L., Lane H., 2002, *Economy of effort in different speaking conditions. I. A preliminary study of intersubject differences and modeling issues*, «The Journal of the Acoustical Society of America», 112, 1627-1641.
- Picheny M.A., Durlach N.I., Braida L.D., 1985, *Speaking Clearly for the Hard of Hearing I: Intelligibility Differences between Clear and Conversational Speech*, «Journal of Speech and Hearing Research», 28, 96-103.
- Picheny M.A., Durlach N.I., Braida L.D., 1986, *Speaking Clearly for the Hard of Hearing II: Acoustic Characteristics of Clear and Conversational Speech*, «Journal of Speech and Hearing Research», 29, 434-446.
- Roach P., 1998, *Some languages are spoken more quickly than others*, in Bauer L., Trudgill V. (a cura di), *Language myths*, London, Penguin, 150-158.
- Schmid S., 2003, *Aspetti prosodici del Foreigner Talk italiano*, in Valentini A., Mulinelli P., Cuzzolin P., Bernini G. (a cura di), Atti del XXXVI congresso internazionale di studi della Società di Linguistica Italiana (SLI) *Ecologia linguistica* Bergamo, 26-28.IX.2002, Roma, Bulzoni, 347-368.
- Schulman R., 1989, *Articulatory dynamics of loud and normal speech*, «The Journal of the Acoustical Society of America», 85, 295-312.
- Smiljanić R., Bradlow A.R., 2005, *Production and perception of clear speech in Croatian and English*, «The Journal of the Acoustical Society of America», 118, 1677-1688.
- Uchanski R.M., 2005, *Clear speech*, in Pisoni D.B., Remez R.E. (a cura di), *The handbook of speech perception*, Oxford & Malden, Mass., Blackwell, 207-235.

- Uchanski R.M., Choi S.S., Braida L.D., Reed C.M., Durlach N.I., 1996, *Speaking clearly for the hard of hearing IV: Further studies of the role of speaking rate*, «Journal of Speech and Hearing Research», 39, 494-509.
- Uther M., Knoll M.A., Burnham D., 2007, *Do you speak E-NG-L-I-SH? A comparison of foreigner- and infant-directed speech*, «Speech Communication», 49, 2-7.
- Wright R., 2003, *Factors of lexical competition in vowel articulation*, in J. Local, R. Ogden (a cura di), *Phonetic interpretation. Papers in Laboratory Phonology*, Temple Rosalind, Cambridge, Cambridge University Press, VI, 75-87.

Informazioni sugli autori

Silvia Calamai
Laboratorio di Linguistica
Scuola Normale Superiore – Pisa.
Email:

calsilvia@tiscali.it

calamai@sns.it