

Sezione V

METODI PER L'INDAGINE FONETICA E PROSODICA

CINZIA AVESANI[○] – MARIO VAYRA[◇] – CLAUDIO ZMARICH[○]
([○]ISTC-CNR, Padova – [◇]Università di Bologna)

Coordinazione vocale-consonante e preminenza accentuale in italiano. La sfida della *Articulatory Phonology**

1. INTRODUZIONE

1.1.

L'obiettivo di questo lavoro è di far luce sui processi articolatori che governano la coordinazione vocale-consonante in italiano e la produzione della preminenza accentuale.

È da lungo tempo patrimonio delle scienze fonetiche la nozione che consonanti e vocali s'influenzano reciprocamente, durante il parlato, nelle rispettive caratteristiche temporali e spettrali (si veda ad esempio il trattato di Edward Search, del 1773, citato in Munhall *et al.*, 1992). Rimane tuttavia alquanto oscuro il modo in cui si realizza la loro coordinazione sequenziale. Un'ottima fonte di informazione è costituita, com'è noto, dalle durate acustiche dei segmenti vocalici e consonantici, o, più propriamente, dagli andamenti delle variazioni delle durate acustiche, misurate calcolando la lunghezza di intervalli compresi fra eventi *convenzionalmente* scelti, quali, per una vocale fra occlusive sorde, ad esempio, il rilascio della chiusura consonantica precedente e la cessazione del processo di sonorizzazione che precede la fase di occlusione dell'occlusiva seguente. La struttura temporale di questi intervalli, così misurati, varia sistematicamente, da lingua a lingua, in funzione di molteplici fattori, sia segmentali, sia prosodici (Lehiste, 1970), quali: il tempo di elocuzione (Gay, 1981), il grado di accento (Lehiste, 1970; Oller, 1973), il tipo più o meno spontaneo o formale di parlato (Lindblom, 1990), la scansione prosodica dell'enunciato (*phrasing*) e il contesto fonetico.

* Il lavoro è stato sviluppato congiuntamente dai tre autori. Tuttavia, a scopi accademici, Cinzia Avesani deve considerarsi responsabile per il paragrafo 4; Cinzia Avesani e Claudio Zmarich per i paragrafi 2.4 e 3; Mario Vayra per i paragrafi 1, 2.1, 2.2, 2.3; i tre autori congiuntamente per il paragrafo 5. Ringraziamo Graziano Tisato per il suo essenziale supporto all'utilizzo di *ELITE*.

Di particolare interesse nell'indagine sui meccanismi soggiacenti la coordinazione V-C si è rivelato il confronto fra i due seguenti fattori di variazione della durata acustica della vocale:

- 1) il contesto consonantico postvocalico, entro la sillaba;
- 2) lo status accentuale della sillaba, ai diversi livelli della gerarchia prosodica: sillaba atona, sillaba tonica con accento lessicale (*stress*), sillaba tonica con accento lessicale e prominente intonativa (*pitch accent*) (cfr. ad es. Beckman, 1986; Nespor e Vogel, 2007; Gussenhoven, 2004).

1.2. “*Compensatory Shortenings*”

Per quanto concerne il primo fattore, sappiamo che ‘quasi universalmente’ le vocali sono più brevi in sillaba chiusa che in sillaba aperta (cfr. Maddieson, 1985, per una rassegna), davanti a nessi consonantici che a consonanti singole (Lindblom *et al.*, 1981; Fowler, 1983), davanti a consonanti sorde che a sonore (Chen, 1970) e davanti a labiali che ad alveolari o velari (cfr. Lehiste, 1970, per una rassegna). Questi fenomeni di accorciamento fanno parte di un’ampia famiglia, cui ci si riferisce in genere con il termine di accorciamenti compensativi (*compensatory shortenings*; cfr. Lindblom e Rapp, 1973). Ciò che essi hanno in comune è che l’aumento di durata di un’unità linguistica dovuto a un certo fattore appare parzialmente controbilanciato da un accorciamento temporale altrove, entro quell’unità (Fowler, 1981; 1983). Nel caso specifico delle consonanti in posizione post-vocalica la generalizzazione sembra essere che un aumento nella durata della consonante in posizione di coda si associa a una diminuzione della durata del nucleo vocalico (cfr. ad esempio, Lehiste, 1970; Lindblom *et al.*, 1981).

Un’interpretazione comunemente accettata di questa generalizzazione è che essa rifletta una tendenza da parte del parlante a comprimere la durata di una vocale al fine di mantenere costante la durata globale della sillaba come unità (cfr. recentemente, Flemming, 2001, entro un quadro fonologico *constraint-based*, vicino per certi aspetti alla *Optimality Theory*). In quanto tali, questi accorciamenti sono stati conseguentemente visti come sintomatici di una tendenza delle lingue a mantenere l’isocronia di certe unità del parlato (Lindblom *et al.*, 1981). Forti dubbi sulla natura ‘compensativa’ di questi accorciamenti, tuttavia, sono stati più volte sollevati da Fowler (1981, 1983). Secondo Fowler, l’accorciamento acustico della vocale non sarebbe che l’effetto, epifenomenico, della sovrapposizione coarticulatoria fra consonanti e vocali, entro un determinato dominio fonetico (su queste linee, cfr. anche Vayra, 1997, per una discussione orientata sull’italiano). Seguendo Öhman (1967) Fowler suggerisce che, dove possibile, la produzione della vocale in un enunciato è ciclica: cioè, i gesti vocalici sono sequenziali e adiacenti, con le consonanti in qualche modo sovrapposte alle voca-

li, come singoli eventi locali. Da questa prospettiva, “vowels constitute a somewhat separate articulatory stream from gestures involved in a consonant production” (Fowler, 1983: 393). Conseguentemente, le vocali dovrebbero avere la ‘responsabilità’ maggiore nel preservare l’ordine seriale di consonanti e vocali.

Per l’inglese, i dati acustici e articolatori presentati in Munhall *et al.* (1992), ove viene esaminato l’accorciamento di una vocale davanti a consonanti tautosillabiche, sembrano confermare questa interpretazione. Munhall e colleghi non trovano alcuna prova di accorciamento articolatorio del gesto di apertura di una vocale quando una o più consonanti sono aggiunte in posizione di coda sillabica. L’abbassamento della mandibola per la vocale ha la stessa durata e la stessa ampiezza indipendentemente dalla lunghezza della coda. Ciò che essi riferiscono, invece, è che l’*onset* del gesto di chiusura della mandibola per le consonanti in posizione di coda viene anticipato.

Per quanto riguarda l’italiano, due autori del presente lavoro (Vayra *et al.*, 1999) hanno fornito indizi acustici a favore dell’origine coarticolatoria dell’accorciamento ‘compensativo’ della vocale in sillaba chiusa. In accordo con i risultati ottenuti per l’inglese da Munhall *et al.* (1992), quel lavoro suggerisce che in italiano l’accorciamento acustico della vocale in sillaba chiusa *vs.* aperta è determinato da un attacco anticipato del gesto di chiusura associato alla coda consonantica. Secondo l’interpretazione ivi proposta, l’anticipazione del gesto di chiusura consonantica causa una *troncamento* del movimento di apertura associato alla vocale. Nell’ipotesi che vocale e consonante siano coprodotte, l’accorciamento acustico della vocale rappresenterebbe pertanto un caso di *blending* articolatorio di due gesti orali (apertura vocalica e chiusura consonantica), che competono in un dato intervallo di tempo per l’uso di un articolatore comune (la mandibola).

La domanda che in questo lavoro ci poniamo è dunque la seguente: corrispondono a ‘reali’ accorciamenti articolatori della vocale gli accorciamenti acustici che così agevolmente è possibile osservare e misurare in un oscillogramma? Per rispondere a questa domanda esamineremo e confronteremo l’organizzazione articolatoria che governa l’allungamento/accorciamento acustico della vocale in sillaba aperta *vs.* chiusa con l’allungamento/accorciamento acustico della vocale in presenza/assenza di prominente accentuale. Esistono infatti processi articolatori di controllo diversi dalla sovrapposizione coarticolatoria, i quali potrebbero dare origine all’accorciamento/allungamento acustico di una vocale. La prominente accentuale, ad esempio, sembra essere prodotta mediante il controllo *diretto* di movimenti articolatori individuali associati alla vocale. Consideriamo pertanto brevemente alcuni correlati acustici e articolatori della prominente.

1.3. Prominente accentuale

Variabili acustiche. Rispetto alle vocali atone, le vocali toniche (accentate lessicalmente) presentano stadi stabili di durata maggiore e sono meno centralizzate nello spazio formantico. Le transizioni verso lo stadio stabile e fuori da esso

sono più ripide per le vocali toniche che per le atone. La frequenza fondamentale è generalmente più alta e mostra andamenti più marcati di quella delle atone e, *ceteris paribus*, le vocali toniche hanno anche una intensità maggiore delle atone (cfr., ad esempio, Lehiste, 1970). Infine, la pendenza del *locus equation* che correla l'*onset* della transizione al centro di F2 nella vocale di una sillaba tonica CV è più piatta della pendenza dell'equazione in una sillaba atona (Krull, 1989; Lindblom *et al.*, 2007), indice, questo, di una minore influenza coarticulatoria della vocale tonica sulla consonante precedente rispetto alla atona. In altre parole, la 'resistenza coarticulatoria' di un segmento accentato è maggiore di quella di un segmento atono.

Variabili articolatorie. L'attività elettromiografica dei muscoli reclutati per produrre consonanti e vocali con accento lessicale è maggiore dell'attività rilevata per le atone. Nella produzione di sillabe accentate, inoltre, l'attività dei muscoli intercostali interni aumenta; c'è dunque un aumento dello sforzo espiratorio nella produzione di sillabe toniche rispetto alle sillabe atone (Ladefoged, 1967). Per quanto riguarda le misure cinematiche, le vocali atone sono prodotte con movimenti che hanno durata minore, mostrano uno spostamento minore (hanno cioè una minore ampiezza) e una minore velocità massima rispetto a sillabe toniche con accento lessicale, prodotte alla stessa velocità di eloquio (Kuhlen e Moll, 1976; Ostry *et al.*, 1983). Inoltre, le vocali basse sono associate a una posizione della mandibola più bassa quando sono accentate rispetto a quando sono atone e il loro stadio stabile è tenuto più a lungo (Summers, 1987).

Per quanto riguarda l'italiano, Vayra e Fowler (1987, 1992) hanno offerto prove acustiche (durata, ampiezza, F0, F1) e articolatorie (movimenti della mandibola) che in italiano, così come in inglese, una vocale atona bassa è più breve rispetto a una tonica e mostra una riduzione nella dimensione dell'altezza (F1) che interessa il *gestoglobale* per la vocale. Questi risultati hanno avuto ampie conferme: in tutte le sillabe atone, indipendentemente dall'altezza della vocale (con l'unica eccezione rappresentata dalla vocale alta posteriore [u]), l'apertura mandibolare per il gesto vocalico risulta minore che nelle sillabe toniche. Anche all'analisi elettropalatografica risulta che l'apertura del tratto vocale è significativamente minore per le sillabe atone rispetto alle toniche, mentre il grado di innalzamento del dorso della lingua tende, in atonia, a ridursi per le vocali alte ([i,u]) e ad aumentare per la vocale bassa ([a]) (Farnetani e Faber, 1992; Magno Caldognetto *et al.*, 1995; Farnetani, 2003). Inoltre, la pendenza del *locus equation* delle sillabe atone è maggiore di quello di sillabe più prominenti (sillabe accentate lessicalmente e sillabe accentate intonativamente), indice, questo, che la resistenza all'influsso coarticulatorio della vocale sulla consonante varia sistematicamente con il grado di prominenza di una sillaba: minore per le atone e progressivamente maggiore per toniche e accentate intonativamente (Zmarich *et al.*, 2006).

Compatibilmente con questi risultati e con una tendenza generale delle lingue alla riduzione delle vocali in posizione atona (Lindblom, 1963), Vayra *et al.* (1999) mostrano che in italiano l'accorciamento dovuto alla deaccentazione implica una riduzione globale del gesto di apertura per la vocale, evidente già dall'inizio del gesto. Le vocali atone sono soggette a riduzione fonetica, almeno nella dimensione dell'altezza (sui fenomeni di riduzione si veda Albano Leoni *et al.*,

1995; per un nuovo metodo di analisi e quantificazione della riduzione vocalica si veda Savy *et al.*, 2005). La riduzione, una caratteristica spesso attribuita alle lingue ad isocronia accentuale (cfr. , ad esempio, Dauer , 1983; Bertinetto, 1981), sembra quindi giocare un ruolo nella fonetica dell'accento in italiano, tradizionalmente descritto come una lingua a isocronia sillabica.

Tutti questi risultati sono compatibili con la proposta che la prominza (accento lessicale e/o prominza intonativa) consista in un aumento globale dello sforzo articolatorio (Öhman, 1967; Lehiste, 1970); e con la scoperta di Kelso *et al.* (1983) che i movimenti delle mani prodotti in sincronia con le sillabe toniche presentano ampiezza maggiore dei movimenti prodotti in sincronia con le sillabe atone, nonostante l'esplicita istruzione impartita ai soggetti dell'esperimento di mantenere l'ampiezza dei movimenti costante.

Risultati nuovi sulla prominza accentuale sono stati presentati da studiosi che applicano all'analisi del parlato la teoria dei *sistemi dinamici*. Un sistema dinamico è un insieme di leggi che specifica le forze che modificano, nel tempo, certi aspetti di un sistema. Modelli dinamici sono stati usati per catturare le caratteristiche di molti fenomeni naturali; fra questi, il controllo del movimento umano specializzato. Nel movimento specializzato, devono essere coordinate molte parti del corpo in modo tale che gli andamenti che emergono siano stabili sul piano spaziale e temporale, e generabili reiteratamente.

Il quadro offerto dalla teoria dei sistemi dinamici è utile perché offre una spiegazione unificata non soltanto delle forme 'superficiali' del movimento e delle loro proprietà di stabilità, ma anche del modo in cui tali forme si alterano quando i parametri 'soggiacenti' del sistema sono modificati (si veda: Turvey, 1977, per l'applicazione di questa prospettiva all'azione, in generale; Saltzman e colleghi (Saltzman, 1995; Saltzman e Kelso, 1987; Saltzman e Munhall, 1989) e Browman e Goldstein (1995) per un modello specifico di produzione del parlato. Per un'applicazione all'accento lessicale si veda, ad esempio, Ostry *et al.* 1983; Kelso *et al.*, 1985; Vatikiotis-Bateson, 1987; su accento e prominza intonativa, si veda Beckman *et al.*, 1992; Beckman *et al.*, 1992; Fowler , 1995; Harrington *et al.*, 1995; Cho, 2006).

2. LA "ARTICULATORY PHONOLOGY" E LA PROSPETTIVA DELLA "TASK DYNAMICS"

Questo lavoro si inquadra all'interno di una particolare teoria fonologica e di un particolare modello del sistema di produzione linguistica. Specificamente, adotteremo la *Articulatory Phonology* come teoria fonologica (Browman e Goldstein, 1986; 1989; 1990; 1992; 1995; 2000) e la *Task Dynamics* come modello quantitativo che implementa le unità fonologiche postulate dalla *Articulatory Phonology* entro il sistema di produzione linguistica (cfr. , ad esempio, Saltzman 1986; Saltzman e Munhall, 1989; Saltzman e Byrd, 2000; cfr. anche Hawkins, 1992).

2.1. La “Articulatory Phonology”

La *Articulatory Phonology* (*AP*) è una teoria fonologica (ma si vedano le obiezioni di Clements (1992) circa lo *status fonologico* della teoria), la quale assume, seguendo sostanzialmente Fowler (1980), che i fallimenti continuamente perpetuatisi nell’identificare entro il segnale acustico unità fonologiche empiricamente osservabili o ‘pubbliche’ (Goldstein e Fowler, 2003) discendano dalla scelta di un *livello di descrizione* troppo superficiale dell’atto di parola, e che sia invece possibile decomporre l’azione del tratto vocale durante la produzione del parlato in unità discrete ricombinabili. Secondo la *AP*, tra struttura fonologica e struttura fisica vige una relazione *naturale*, potenzialmente governata da leggi. Più specificamente, la tesi è che sia possibile concepire le due strutture come descrizioni, l’una ‘macroscopica’ (*low-dimensional*), l’altra ‘microscopica’ (*high-dimensional*), di un singolo sistema auto-organizzato complesso (per una discussione su questo punto, si veda Vayra, 2003).

La *AP* ipotizza che una descrizione del parlato in termini di *gesti articolatori* offra strumenti che catturano proprietà sia fisiche, sia fonologiche del segnale vocale. Entro questo quadro teorico, il flusso parlato è decomposto in una serie di gesti che controllano le azioni di insiemi distinti di articolatori nel tratto vocale. Questi gesti sono simultaneamente *unità d’azione* e *unità di informazione*.

I gesti sono unità d’azione in quanto costituiscono delle *strutture coordinative* (cfr. Turvey, 1977; Fowler, 1980). Ogni gesto è rappresentato come un sistema dinamico neuromotorio, che guida la coordinazione di un insieme di articolatori e muscoli (potenzialmente indipendenti) nella formazione di una specifica costrizione del tratto vocale. Un gesto di chiusura labiale, ad esempio, comporta l’azione coordinata di tre articolatori: labbro superiore, labbro inferiore e mandibola, insieme a una quindicina di muscoli distinti. Quando un’occlusione labiale è attiva nel tratto vocale, tutti questi articolatori e muscoli sono funzionalmente vincolati, in modo tale che le loro attività cooperino al raggiungimento di una ‘meta’ (*task*): l’occlusione labiale.

I gesti costituiscono delle unità d’informazione, in quanto potenzialmente *distintivi*. Gesti individuali possono essere usati per distinguere il significato di due parole, cioè per creare un’opposizione fonologica, mediante ad esempio la presenza vs. assenza di un gesto particolare, o mediante differenze relative a luogo o grado di una medesima costrizione gestuale.

In sintesi, i gesti sono unità astratte d’opposizione fonologica (una proprietà ‘macroscopica’), ma, essendo riconducibili a sistemi dinamici che guidano la coordinazione di molteplici articolatori e muscoli, essi sono intrinsecamente definiti in relazione alle loro proprietà fisiche (‘microscopiche’).

Questa visione della relazione fra struttura fonologica e produzione linguistica postula che non vi sia mediazione (né necessità di mediazione) fra rappresentazione fonologica e implementazione fonetica. Le unità d’azione (fonetiche) e le unità di informazione (fonologiche) costituiscono una e una sola cosa, e la rappresentazione include tutto ciò che è necessario per l’esecuzione.

La *AP* fa tre ipotesi intorno alla natura delle unità fonologiche, che consentono loro di assolvere al doppio ruolo di unità d’azione (continue e dipendenti dal

contesto) e unità combinatorie e distintive (discrete e indipendenti dal contesto). Queste ipotesi concernono:

- i) l'analisi dell'attività del tratto vocale in azioni di costrizione compiute da *organi* vocali distinti, sostanzialmente corrispondenti agli articolatori 'attivi' della moderna fonologia. Nella sua attuale configurazione computazionale il modello, disegnato sull'inglese, prevede: labbra, apice (*tip*), dorso e radice della lingua, velo e laringe (cfr Halle, 1983). Tali azioni vengono denominate *gesti articolatori*;
- ii) la coordinazione delle azioni in strutture temporaneamente sovrapposte (*gestural constellations*);
- iii) la possibilità di usare la dinamica per costruire un modello computazionale della formazione delle costrizioni nel tratto vocale.

Da un punto di vista sperimentale, il presente lavoro investiga specificamente gli ultimi due punti. (Per una più dettagliata rassegna di questioni attinenti alla *Articulatory Phonology* si rinvia a Vayra, 2003).

2.2. Coordinazione dei gesti e sovrapposizione

Mentre le teorie fonologiche tradizionali assumono che le unità primitive si combinano formando sequenze lineari, la *AP* avanza l'ipotesi che i gesti siano coordinati in strutture di tipo molecolare più elaborate ('costellazioni'), entro le quali essi possono sovrapporsi temporalmente.

Nella prospettiva della *AP*, tale *coproduzione* di gesti articolatori adiacenti è la chiave che può spiegare gran parte della variabilità contestuale osservabile nella lingua parlata. In accordo alla *AP*, la ragione di ciò risiede nella natura dei distinti 'organi' di costrizione, i quali condividono articolatori e muscoli. Quando due gesti si sovrappongono, infatti, le attività dei gradi di libertà meccanici associati ai singoli gesti dipendono da entrambi i gesti.

Consideriamo ad esempio la coproduzione, in italiano, di un gesto consonantico di costrizione laminale con il gesto dorsale associato a vocali diverse, come in [di] e [du]. Adattando all'italiano il modello della *AP* si può ipotizzare che in entrambi i casi venga prodotto il medesimo gesto apicale (indipendente dal contesto), ma che i contributi dei vari gradi di libertà (apice e dorso della lingua) siano diversi perché diverse sono le restrizioni imposte dal gesto vocalico¹. La

¹ Nell'attuale implementazione computazionale della *AP*, l'apice della lingua (*tongue tip*) costituisce l'*end effector*, ovvero l'articolatore terminale, di una catena di articolatori, reciprocamente vincolati in una struttura coordinativa finalizzata alla formazione di una costrizione: apice della lingua, dorso della lingua e mandibola. L'importante distinzione fra suoni coronali

teoria della *Task Dynamics* fornisce un modello formale di tale variabilità contestuale (cfr. Saltzman, 1986; 1995; cfr. § 2.3).

Nella *AP*, i gesti si combinano in ‘molecole’ più ampie grazie alla coordinazione reciproca (*bonding*) dei singoli gesti. Questa coordinazione si compie, entro un modello dinamico, specificando la *relazione di fase* tra gli oscillatori accoppiati che controllano la produzione dei gesti coordinati (Browman e Goldstein, 1990; Saltzman e Byrd, 2000; Goldstein e Fowler, 2003).

Nel suo attuale assetto computazionale la *AP* produce, per un qualsivoglia enunciato (in inglese), una ‘struttura gestuale’ o *gestural score* che consiste di un insieme di gesti e della specificazione del modo in cui questi sono temporalmente coordinati fra loro. Vari lavori di Browman e Goldstein sono volti a mostrare come tali strutture gestuali possano catturare generalizzazioni soggiacenti a numerosi tipi di alternanze fonetiche e fonologiche.

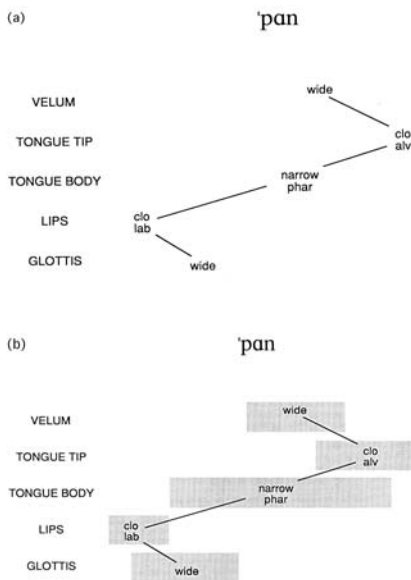
La Figura 1 (tratta da Browman e Goldstein, 1995) mostra la ‘struttura gestuale’ della parola *pawn* con la vocale posteriore non arrotondata, comune nella pronuncia dell’inglese americano ([‘p an]). Le righe orizzontali rappresentano i ‘livelli’ (*tiers*) associati a ciascuno dei distinti ‘or gani’ attualmente utilizzati dal modello. Le etichette associate a ciascun gesto (‘clo’, ‘alv’, ad esempio) sono abbreviazioni per i valori numerici (o campi di valori numerici) dei parametri dinamici che specificano la ‘meta’ (*task o tract variable*) di una determinata costrizione. Così, il gesto apicale (*Tongue Tip*) alla fine della parola *pawn* è specificato, relativamente al ‘grado di costrizione’, per un valore che produrrà chiusura completa, e, relativamente al ‘luogo di costrizione’, per un valore che situerà la restrizione all’altezza della cresta alveolare.

Le linee collegano le particolari coppie di gesti, entro la parola, la cui coordinazione è specificata *esplicitamente* entro il modello della struttura gestuale (Browman e Goldstein, 1992). La coordinazione dei gesti è ottenuta mediante coordinazione di fase (*phasing*): nella struttura gestuale è cioè indicato che una determinata fase del moto di un membro della coppia coincide con una determinata fase dell’altro membro.

Date le specificazioni di fase e la specificazione dei parametri dinamici intrinseci, il modello calcola gli intervalli di attivazione relativi a ciascun gesto. Il risultato è ciò che è stato denominato uno ‘spartito gestuale’ (*gestural score*). Osserviamo che gli intervalli d’attivazione dei gesti si sovrappongono parzialmente (Figura 1b). Mentre le strutture gestuali sono proprietà fisse, lessicali, di una parola, i valori quantitativi dei parametri dinamici dei gesti e le relazioni di fase possono essere *scalari*, possono cioè essere quantitativamente scalati in funzione di varie condizioni intrinseche all’atto di parola. Lo spartito gestuale riflette questa

apicali quali la apico-alveolare [t] dell’inglese “time” e suoni laminali quali la lamino-dentale [t̪] dell’italiano “timo”, è trattata, in *AP*, in termini di orientamento (angolo) dell’apice della lingua sul piano sagittale in rapporto agli assi di “Luogo” e “Grado di Costrizione”.

Figura 1. (a) 'Struttura gestuale' per la parola "pawⁿ". Le linee collegano i gesti la cui coordinazione è specificata lessicalmente nel modello computazionale. (b) 'Spartito gestuale' (Gestural score). L'estensione orizzontale dei riquadri rettangolari rappresenta gli intervalli d'attivazione dei gesti (adattata da Browman e Goldstein, 1995)



operazione di *scaling*. Il *gestural score* costituisce l'ingresso al *Task-Dynamic Model* di Saltzman e colleghi (cfr. Saltzman, 1986; 1995; Saltzman e Kelso, 1987). Questo modello calcola la risposta di un gruppo di articolatori, simulati computazionalmente, ai parametri dinamici che specificano una determinata *task variable* (ad esempio: 'tongue tip alveolar closure'). Gli articolatori si muovono in modo coordinato per raggiungere la 'meta' della costrizione.

2.3. La specificazione dinamica

Un aspetto della descrizione fisica del parlato sembra particolarmente in contrasto con una descrizione del parlato in termini di unità discrete: il fatto che gli articolatori si muovono nel tempo in modo continuo. Questo moto continuo viene modellato nella maggior parte delle teorie fonologiche o delle teorie della produzione assumendo che esistano dei *target* per le unità fonologiche e che l'apparato di produzione interpoli questi *target* (ad esempio, Keating, 1990; Pierrehumbert, 1990; Terrier

et al., 1996). Secondo questa prospettiva, “the phonological information (the target) is not part of the production itself – it is hidden within the mind of the producer, and is only communicated through the resulting interpolation” (Goldstein e Fowler, 2003: 7).

Nella *AP* ogni gesto è modellato come un sistema dinamico massa-molla (un oscillatore) che caratterizza la formazione e il rilascio di una costrizione locale entro il tratto vocale (il ‘compito’ o ‘meta funzionale’) del gesto. Un sistema dinamico è un’equazione (o una serie di equazioni) che esprime il modo in cui lo stato di un sistema cambia nel tempo.

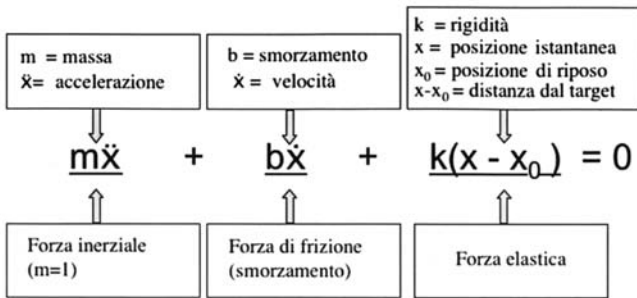
Ad esempio, la sillaba [ba] inizia con un gesto la cui ‘meta’ è la chiusura labiale. La formazione di tale costrizione implica un cambiamento nel tempo della distanza fra labbro superiore e inferiore. Tale cambiamento è modellato usando un sistema dinamico di secondo ordine (*point attractor*), che regola la formazione di una costrizione per mezzo dell’organico pertinente. La specificazione dinamico-funzionale di un gesto include i seguenti parametri: *posizione di equilibrio*, *rigidità* (*stiffness*) e *smorzamento*.

La posizione di equilibrio del sistema dinamico corrisponde al *target* fonetico da conseguire (*tract variable*). Durante il periodo di attivazione di un gesto *labiale*, ad esempio, al parametro ‘posizione di equilibrio’ è assegnato il valore relativo alla chiusura labiale. Quel valore consente la realizzazione della ‘meta’ fonetica della chiusura labiale. La rigidità è correlata alla costante temporale della formazione della costrizione. Lo smorzamento si assume generalmente sia di tipo critico, in modo tale che il sistema possa raggiungere la propria posizione d’equilibrio, senza oltrepassarla (il che rappresenterebbe un caso di *overshooting* del *target*).

I valori parametrici di rigidità e smorzamento, combinati, determinano la quantità di tempo che il sistema impiega per conseguire la meta.

Occorre notare qui due punti di capitale importanza per la teoria e l’analisi fonetica e fonologica. Primo, un gesto può essere modellato come un sistema dinamico con un numero fisso di valori parametrici. Ciò significa che anche se gli articolatori organizzati in un gesto (ad esempio, labbro inferiore, labbro superiore e mandibola nel gesto di Apertura Labiale) si muovono in modo continuo, l’equazione che descrive il gesto è fissa e invariante nel tempo. La proprietà fondamentale di una descrizione dinamica è dunque che essa non contiene in sé la rappresentazione della durata del movimento, e tuttavia caratterizza i movimenti articolatori nello spazio e nel tempo, come funzione dei parametri del sistema dinamico. Secondo, il *modo* in cui gli articolatori si muovono nel tempo è specifico del particolare sistema dinamico implicato, e dunque l’identità della unità gestuale invariante risulta specificata *direttamente* dal movimento stesso.

Esaminiamo più dettagliatamente l’equazione che descrive la dinamica di un sistema massa-molla. Essa include componenti di forza (inerziale, di frizione ed elastica) associate a 4 parametri:



x_0 rappresenta il *target* del gesto; specifica la posizione della massa quando la molla è alla sua lunghezza di riposo o posizione di equilibrio;

m è la massa; si assume che essa sia uguale ad 1 per tutti i gesti;

k è il parametro rigidità, e rappresenta l'elasticità del sistema; è proporzionale alla 'velocità' intrinseca del gesto. A parità di condizioni, un gesto con k minore avrà una durata intrinsecamente maggiore di un gesto con k più alta.

Infine, b è il parametro dello smorzamento.

Poiché nel modello la massa ha sempre il valore di 1 e il rapporto di smorzamento ($b/(2[mk]^{1/2})$) è impostato a 1, per risolvere l'equazione del movimento nel tempo di una *tract variable* dobbiamo specificare i valori relativi alla posizione di riposo x_0 e alla rigidità k .

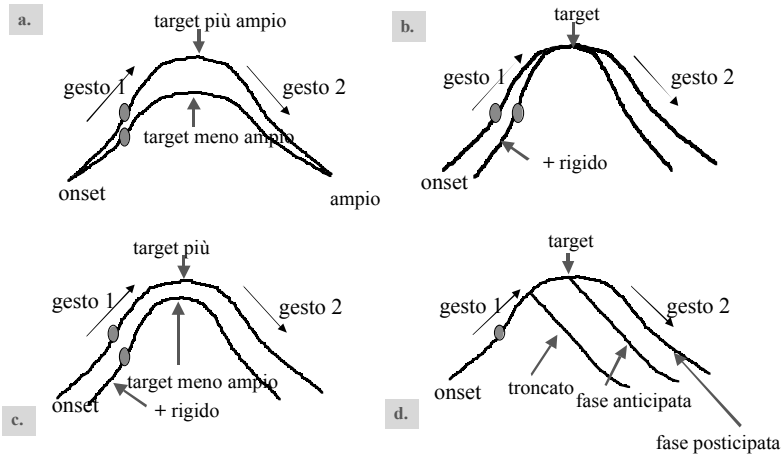
2.4. Conseguenze cinematiche della manipolazione dei parametri dinamici nel modello della "Task Dynamics"

Le caratteristiche dei movimenti articolatori che risultano dall'esecuzione dei gesti dipendono dunque dai valori dei parametri che vengono specificati per un certo gesto. I parametri rilevanti sono il *target*, che rappresenta l'ampiezza soggiacente del movimento, la rigidità, che rappresenta la frequenza di oscillazione naturale del sistema, il rapporto di smorzamento, la fase tra gesti contigui e il loro tempo di attivazione. È quindi un punto cruciale del modello il fatto che tutte le variazioni sistematiche articolatorie o cinematiche possano essere interpretate come conseguenze dell'impostazione dei parametri dinamici².

² Secondo Byrd e Saltzman (2003) le caratteristiche cinematiche di un gesto di costrizione, in particolare le sue proprietà temporali, possono essere ulteriormente modificate dalla concomitante attivazione di un gesto prosodico o π -gesture. I π -gesture non hanno una realizzazione articolatoria indipendente: si estendono su un arco temporale definito e si realizzano in modo vicario attraverso gli effetti che ingenerano sulle proprietà dinamiche di tutti i gesti di costrizione che siano attivi contemporaneamente. Gli autori hanno applicato la nozione di gesto prosodico per spiegare e simulare gli effetti temporali di allungamento finale e allungamento/raf forzamento iniziale indotti su gesti di costrizione dalla presenza e dalla diversa forza di un confine

Byrd *et al.*, 2000 e Cho (2006) hanno fornito utili schemi riassuntivi delle conseguenze cinematiche delle diverse manipolazioni dei parametri nell'equazione di moto (1). Li rappresentiamo nella Figura 2, che visualizza le manifestazioni cinematiche idealizzate che conseguono a quattro diverse specificazioni dei parametri dinamici.

Figura 2. *Conseguenze cinematiche idealizzate delle diverse specificazioni dei parametri dinamici: a) cambiamento nel target; b) cambiamento nella rigidità; c) scalatura lineare; d) cambiamento di fase (adattata da Cho, 2006)*



a) *cambiamento nel target* (cfr. Figura 2a): il *target* corrisponde all'ampiezza articolatoria del gesto. Un cambiamento nella specificazione del *target* induce un cambiamento nella quantità di spostamento compiuto dagli articolatori reclutati in un gesto, senza alterare la durata dell'intervallo temporale che intercorre dall'*onset* del movimento al raggiungimento del *target*. Se nell'equazione di moto cambiamo solo i valori del parametro x_0 (un cambiamento di *target* puro) la velocità massima del gesto e lo spostamento cambiano in modo direttamente proporzionale senza che ciò comporti un cambiamento nella durata del movimento (dall'*onset* al raggiungimento del *target*). Questo perché se il *target* è più distante (e la rigidità è costante), gli

prosodico. Il livello di attivazione di un π -gesture è determinato dalla forza del confine prosodico, e gli effetti generati sono proporzionali al livello di attivazione. In questa sede non discuteremo la possibile estensione della nozione di gesto prosodico all'interpretazione delle variazioni cinematiche indotte dalla prominenza accentuale.

articolatori devono compiere un tragitto maggiore nello stesso intervallo di tempo. L'unico modo di raggiungere un *target* più distante senza aumentare la durata dello spostamento è di aumentare la velocità in modo proporzionale all'aumento dello spostamento. In questo caso, come si nota nella Figura 2a, la durata della fase di accelerazione (o *Time-To-Peak velocity*, definita come l'intervallo temporale incluso tra l'*onset* del movimento e il momento in cui viene raggiunta la massima velocità, rappresentata nella Figura 2 come un pallino) rimane costante.

Queste proprietà cinematiche sono le conseguenze naturali della dinamica che controlla il gesto.

b) *cambiamento nella rigidità* (cfr. Figura 2b): il parametro della rigidità, k , è molto importante perché è il parametro che controlla le variazioni nella durata di un movimento: più rigida è la molla, più veloce è il movimento. Se la rigidità è il solo parametro che sottende le differenze cinematiche che osserviamo in un articolatore ci si deve aspettare un cambiamento nella velocità massima che un articolatore raggiunge in un gesto, ma non nell'ampiezza del suo spostamento. Cioè, un gesto con minore rigidità impiegherà più tempo a raggiungere il *target*. La durata dell'accelerazione (*Time-To-Peak velocity*, o *TTP*) covaria in modo inversamente proporzionale alla rigidità. Gestii con una rigidità minore raggiungeranno la loro massima velocità più tardi rispetto a gesti con una rigidità maggiore.

c) *Cambiamenti proporzionali di target e rigidità: scalatura lineare* (cfr. Figura 2c): ci sono altri mezzi con i quali si possono generare variazioni di durata dei gesti entro un sistema massa-molla. Uno di questi è la scalatura lineare (Harrington *et al.*, 1995). Possiamo definirla come un cambiamento proporzionale uniforme, nello spazio e nel tempo, della cinematica di un gesto; nella scalatura lineare a un cambiamento nell'ampiezza di un gesto – e quindi nei valori del suo *target* – si abbina un cambiamento proporzionale nella sua rigidità (più ampio il *target*, minore la rigidità). Si può pensare alla scalatura lineare come ad un unico parametro dinamico che può essere alla base delle variazioni nei gesti condizionate dalla prosodia (Harrington *et al.*, 1995; Byrd *et al.*, 2000). La scalatura che contrae un gesto spazialmente e temporalmente produrrà gesti con le seguenti caratteristiche cinematiche: minore durata, minore spostamento, minore *TTP*, ma uguale velocità massima.

d) *Cambiamenti nella sincronizzazione intergestuale (fase): troncamento ed estensione ('de-troncamento')* (cfr. Figura 2d): con troncamento ed estensione o 'de-troncamento' (cfr. Harrington *et al.*, 1995; Byrd *et al.*, 2000) ci riferiamo alle conseguenze cinematiche indotte dallo spostamento nel tempo dell'attivazione di un gesto rispetto ad un altro. Se due gesti contigui condividono gli stessi articolatori (ad esempio, le labbra) quando l'attivazione del gesto seguente viene anticipata e si sovrappone alla parte finale del gesto precedente, la traiettoria cinematica del gesto precedente sarà troncata, il gesto non potrà raggiungere il suo *target* e

la sua durata diminuisce. Possiamo immaginare che l'attivazione del gesto seguente 'scivoli' indietro nel tempo a coprire la porzione finale dell'intervallo di attivazione del gesto precedente. Specularmente, se l'attivazione del secondo gesto viene ritardata, cioè è 'scivola' in avanti nel tempo, il primo gesto verrà a 'esteso', ovvero sarà realizzato in modo più pieno, sarà più lungo e, in certi casi più ampio.

In un puro cambiamento di sincronizzazione tra gesti non ci si aspetta un cambiamento nella specificazione della velocità massima del gesto né un cambiamento nella specificazione del suo *target* o della sua rigidità, perché il cambiamento nella durata del gesto è una conseguenza diretta della sovrapposizione o scivolamento temporale di un gesto sull'altro. Ma, come specificano Byrd e Salzman (2003), il troncamento può avere conseguenze cinematiche diverse a seconda della forma globale che assume la traiettoria del gesto: se il gesto forma un *plateau*, ovvero se raggiunge il *target* e lo mantiene per un breve intervallo di tempo, e il troncamento si applica principalmente in tale regione, il cambiamento nello spostamento sarà vicino allo zero (si veda la freccia che indica 'fase anticipata' nella Figura 2d); se il gesto non forma un *plateau*, ovvero se il movimento di apertura raggiunge il *target* e non lo mantiene ma immediatamente inizia il movimento di chiusura, e il troncamento si applica prima di tale regione, il cambiamento nello spostamento sarà grande (si veda la freccia che indica 'troncato' nella Figura 2d). Infine, la durata minore (nel caso del troncamento) o maggiore (nel caso dell'estensione) del gesto proverrà dal cambiamento nella durata della fase di decelerazione (della porzione del gesto che va dal punto di massima velocità al *target*), mentre non ci sarà differenza nella durata della fase di accelerazione (*TTP*).

3. L'ESPERIMENTO

3.1. Metodo

Soggetti e corpus. Abbiamo analizzato le caratteristiche cinematiche del gesto di apertura labiale di sillabe CVC CVC prodotte da due soggetti adulti di sesso femminile: SG e CA. Il nostro *corpus* è costituito di sillabe aperte e chiuse caratterizzate da tre livelli di prominenza: nessun accento (sillabe atone), accento lessicale (sillabe toniche), accento lessicale e intonativo (sillabe toniche in *focus* contrastivo; da ora in avanti indicheremo queste sillabe come 'focalizzate'). Le sillabe sono inserite in posizione mediana in pseudo-parole trisillabiche di tipo /CV.CV(C).CV/ nelle quali C è sempre una consonante bilabiale ([b, m]) e V è una vocale massimamente bassa o massimamente alta ([a, i]). La scelta di una consonante occlusiva bilabiale è una scelta indotta necessariamente dal tipo di sistema di acquisizione di dati articolatori utilizzato per questa ricerca (*ELITE*, vedi

infra), con il quale siamo in grado di registrare solamente il movimento degli articolatori visibili: dunque, labbra e mandibola. La scelta della sonorità è legata alla necessità di distinguere acusticamente la fase di occlusione della consonante dall'eventuale presenza di pause. La scelta di una vocale bassa ([a]) permette di massimizzare il contrasto sintagmatico tra la massima costrizione che caratterizza la fase di tenuta articolatoria del fono bilabiale e la massima apertura che caratterizza la vocale. I risultati relativi a queste condizioni che ottimizzano il contrasto sulla dimensione verticale del gesto di apertura labiale potranno essere verificati in una condizione in cui il contrasto sulla dimensione verticale del gesto è invece minimo, quando alla consonante bilabiale segue la vocale alta [i].

Le pseudo-parole sono parte di frasi naturali inserite in brevi brani dialogici. Dopo una fase di familiarizzazione con il *corpus*, il soggetto pronuncia la frase senza leggerla in risposta ad uno degli sperimentatori che legge la propria parte del dialogo. Le frasi sono state ripetute 8 volte, sia a velocità di eloquio normale sia a velocità sostenuta; esclusi i casi che sono risultati non analizzabili, i dati su cui sono basati i nostri risultati ammontano a 5-7 ripetizioni per ciascuna velocità di eloquio.

I fattori considerati nell'esperimento sono dunque i seguenti:

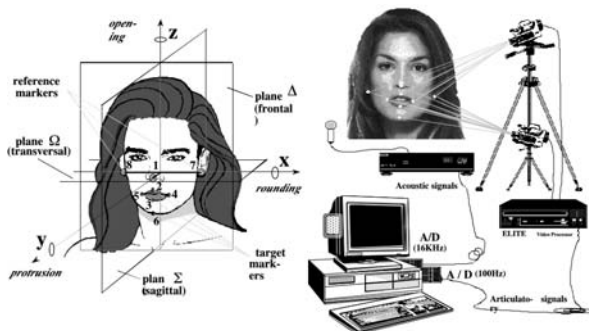
- Stato accentuale della sillaba: atona vs. tonica vs. focalizzata
- Struttura della sillaba: CV vs. CVC
- Tipo di vocale /a/, /i/
- Velocità di eloquio: normale vs. veloce

Strumentazione. Per registrare i movimenti degli articolatori visibili (labbra, mandibola) ci siamo avvalsi del dispositivo *ELITE* (*Elaboratore di Immagini Televisive*, cfr. Ferrigno e Pedotti, 1985; Ferrigno et al., 1999). Questo sistema utilizza la luce infrarossa riflessa da minuscoli *marcatori* passivi incollati sul viso del soggetto. I marcatori riflettono la luce emessa da una corona di *LED* montati attorno all'obiettivo di due telecamere; le telecamere acquisiscono il segnale riflesso dai marcatori grazie ai loro sensori optoelettronici. L'informazione relativa al movimento, acquisita con una frequenza di campionamento di 100 Hz, viene inviata ad un elaboratore di immagini per il riconoscimento dei marcatori e il calcolo del loro baricentro e quindi ad un calcolatore per le elaborazioni successive. Le coordinate dei marcatori vengono ricostruite tridimensionalmente con un grado di precisione che tollera un errore massimo di 0,08 mm per un cubo di 200 mm di lato (vedi Figura 3, parte sinistra).

I marcatori passivi sono posizionati sulla punta del naso, sul punto centrale del labbro superiore (*LS*) e del labbro inferiore (*LI*), sull'angolo sinistro e sull'angolo destro della bocca, sul punto centrale del mento, sui lobi delle orecchie (vedi Figura 3, parte destra). Gli ultimi due servono come punti di riferimento per eliminare gli effetti di eventuali movimenti della testa.

Il sistema calcola la traiettoria del movimento (nella sua dimensione verticale, orizzontale e antero-posteriore), la sua velocità e l'accelerazione istanta-

Figura 3. A sinistra: marcatori e piani 3D di riferimento per la misurazione dei movimenti; a destra: la strumentazione di ELITE



nea. Nello specifico, abbiamo calcolato le traiettorie del *LS* (la distanza ortogonale del marcatore sul labbro superiore dal piano orizzontale virtuale di riferimento che passa per la punta del naso e i lobi delle orecchie), del *LI* (la distanza ortogonale del marcatore sul labbro inferiore dal piano orizzontale virtuale di riferimento che passa per la punta del naso e i lobi delle orecchie), della mandibola (la distanza ortogonale del marcatore sulla mandibola dal piano orizzontale virtuale di riferimento che passa per la punta del naso e i lobi delle orecchie), e dell'apertura labiale (*AL*, definita come la distanza tra i marcatori posti al centro delle labbra).

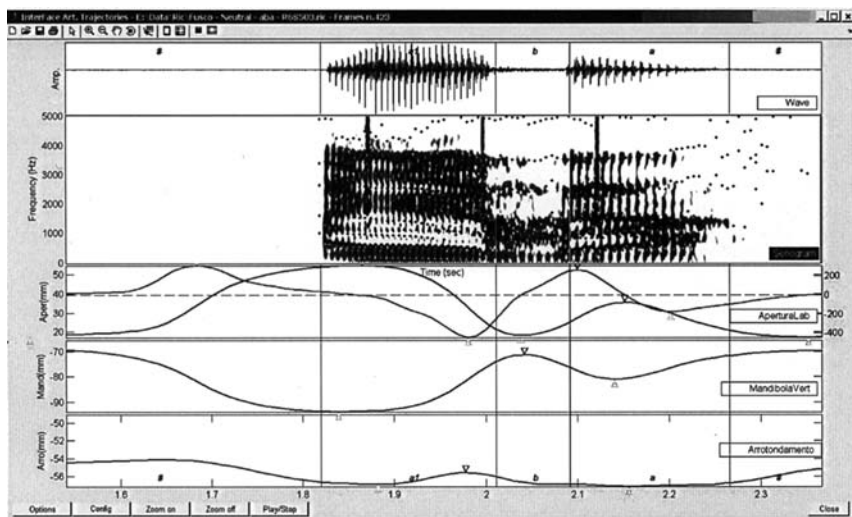
Contemporaneamente alla registrazione cinematografica dei movimenti labiali e mandibolari, viene registrato anche il segnale sonoro su due dispositivi indipendenti: direttamente su *ELITE* (a 16 kHz e 12 bit) e su un registratore *DAT* (a 44 KHz e 16 bit).

Tutte le elaborazioni cinematiche sono state eseguite impiegando il software di gestione di *ELITE* ed esportando i file acustici e cinematografici in uscita (ricostruzioni dell'andamento cronologico filtrato di una serie di parametri articolatori) in *Interface*, un software interattivo creato per l'elaborazione e l'analisi dei dati cinematografici (cfr. Tisato *et al.*, 2005). Nella Figura 4 sono rappresentati i dati acustici e cinematografici sincronizzati della parola "aba" così come vengono rappresentati da *Interface*.

A partire dall'alto, il programma visualizza la forma d'onda, lo spettrogramma, l'andamento della apertura labiale con la relativa curva della velocità, l'andamento della mandibola e dell'arrotondamento delle labbra. I triangoli sulle curve indicano i minimi e i massimi del movimento e della velocità.

Analisi. Le analisi condotte sono di tipo acustico e di tipo articolatorio. Per quanto riguarda le analisi acustiche, ciascuna frase è stata segmentata con *Praat* (www.fon.hum.uva.nl/praat) per misurare la durata della sillaba e della vocale

Figura 4. Rappresentazione acustica e cinematica della parola "aba" in Interface



target, e per valutare la effettiva realizzazione di un accento intonativo nelle frasi prodotte con la focalizzazione della parola *target*. L'intonazione delle parole focalizzate è stata trascritta col sistema *ToBI*.

Per quanto riguarda le analisi articolatorie, delle curve cinematiche relative ai diversi articolatori registrati, abbiamo analizzato in questo studio le proprietà dell'Apertura Labiale (*AL*) relative al ciclo articolatorio della sillaba CV(C), in posizione centrale di parola, nelle tre condizioni di accento considerate: atona, tonica e focalizzata. In particolare, riportiamo in questa sede i dati relativi al gesto di apertura, cioè alla porzione del segnale cinematico compresa tra un punto di minimo spostamento e un punto di massimo spostamento: dal massimo di chiusura di *AL* per la consonante al massimo di apertura di *AL* per la vocale (cfr. Figura 5).

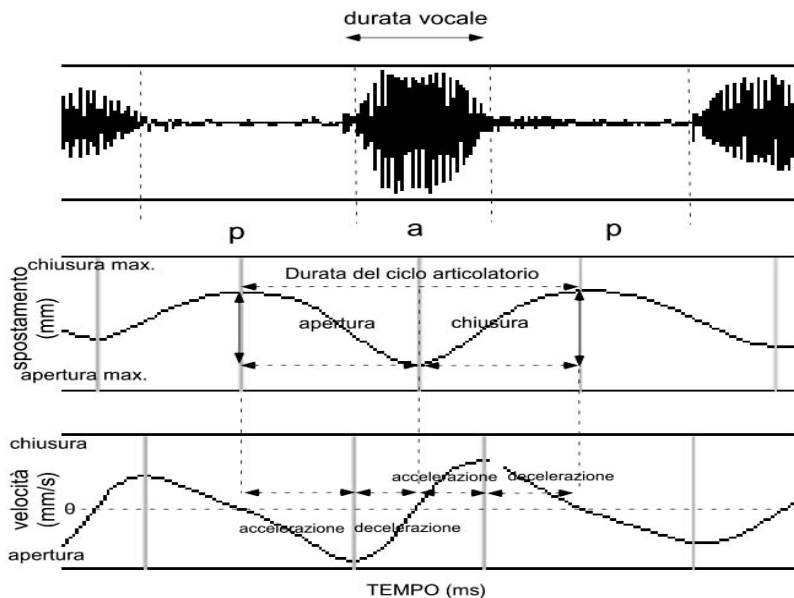
Del gesto di apertura sono stati misurati i valori temporali (in ms) e spaziali (in mm) dei punti di massimo e minimo spostamento e i valori della velocità. A partire da questi, abbiamo calcolato una serie di variabili cinematiche relative al gesto di apertura:

- durata (cfr. Figura 6a)
- ampiezza (cfr. Figura 6b)
- velocità massima (cfr. Figura 6c)
- *TTP – Time-to-Peak velocity*: l'intervallo temporale che intercorre tra l'inizio del gesto ed il punto in cui esso raggiunge la velocità massima. Si tratta di una misura che esprime l'intervallo di accelerazione del gesto (cfr. Figura 6d)

- *TTP%* – *Time-to-Peak velocity* relativo: l'intervallo temporale che intercorre tra l'inizio del gesto ed il punto in cui esso raggiunge la velocità massima, espresso in percentuale rispetto alla durata totale del gesto. Questa variabile descrive implicitamente la proporzione della fase di accelerazione sulla fase di decelerazione: ad esempio, un valore superiore al 50% indica una fase di accelerazione relativamente lunga, cioè un gesto che nella sua fase iniziale è più lento (cfr. Figura 6e).

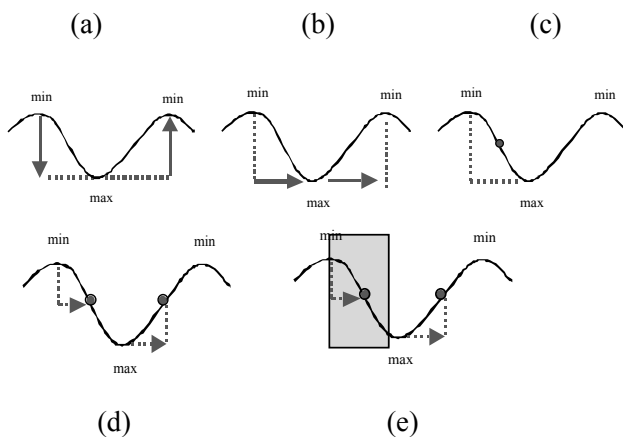
Abbiamo considerato anche una variabile dinamica, *larigidità* del gesto, definita dal rapporto tra velocità massima e ampiezza (velocità massima/ampiezza).

Figura 5. *Illustrazione delle fasi di apertura e chiusura di un ciclo articolatorio. Forma d'onda, curva dello spostamento e curva della velocità*



I dati che discuteremo nel resto dell'articolo sono relativi alla velocità di eloquio normale, alla sillaba CV con vocale [a] e alla sillaba CVC con vocale [a] e [i]. I *test* statistici utilizzati (analisi della varianza a una via, con fattori: grado di prominza della sillaba (atona vs. tonica vs. focalizzata) e struttura della sillaba (aperta vs. chiusa)) sono stati condotti separatamente per ciascun locutore. Presenteremo per primi i risultati cinematici di AL relativi alle diverse condizioni, o gradi, di prominza accentuale; quindi presenteremo i risultati relativi alle diverse strutture sillabiche (sillabe aperte e chiuse); infine, interpre-

Figura 6. Variabili cinematiche usate nell'esperimento: (a) durata; (b) ampiezza/postamento; (c) velocità massima; (d) TTP; (e) TTP%



teremo le variazioni cinematiche in termini del modello proposto dalla *Task Dynamics*.

3.2. Risultati

3.2.1. Gradi accentuali. Sillaba aperta

3.2.1.1. Durata acustica e articolatoria

I dati nella Figura 7 mostrano che per entrambi i soggetti la durata acustica della sillaba aumenta progressivamente in relazione al grado di prominza. Tale differenza è altamente significativa (SG: $F(2,15)=31,126$ $p<0,0001$; CA: $F(2,10)=26,109$ $p=0,0001$). Il *test post hoc* (Fisher) conferma che le sillabe lessicalmente accentate sono più lunghe delle sillabe atone (dato ampiamente noto in letteratura) e più brevi delle focalizzate (cfr. Figura 7).

I dati cinematici relativi alla durata del gesto di apertura, dalla massima chiusura per C alla massima apertura per V, covariano positivamente con la durata acustica, per entrambi i soggetti. Il gesto di apertura è significativamente diverso per i tre gradi di prominza accentuale (SG: $F(2,15)=22,052$ $p<0,0001$; CA: $F(2,10)=23,310$ $p=0,0002$): la durata del gesto di apertura per le atone è significativamente più breve di quella delle toniche e quella delle toniche significativamente più breve di quella delle focalizzate (cfr. Figura 8).

Figura 7. *Durata acustica della sillaba. Locutori CA e SG. Atona <*** tonica <*** focalizzata*

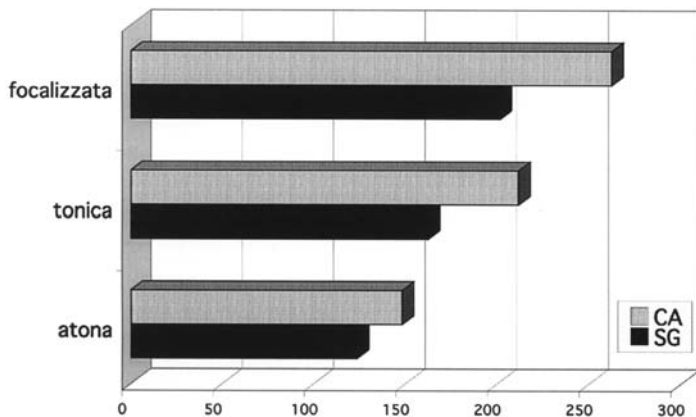
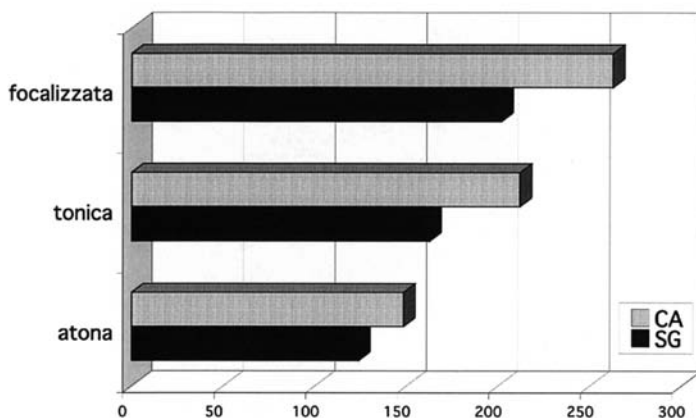


Figura 8. *Apertura labiale. Durata articolatoria del gesto di apertura (dalla massima costrizione per C alla massima apertura per V). Atona <*** tonica <*** focalizzata*



3.2.1.2. Ampiezza e velocità del gesto

Lo stesso andamento si riscontra nei dati relativi all'ampiezza del gesto: l'ampiezza dell'apertura labiale è significativamente diversa nelle sillabe con diverso grado accentuale (SG: $F(2,17)=22,885$ $p<0,0001$; CA: $F(2,10)=31,542$

$p < 0,0001$): per entrambi i soggetti il *test post hoc* (Fisher) conferma che la differenza è altamente significativa per tutti i confronti a coppie (l'ampiezza dell'apertura delle atone è minore di quella delle toniche, e questa è minore di quella delle focalizzate; cfr. Figure 9 e 10).

Dunque, sillabe con gradi progressivamente maggiori di prominenzza sono caratterizzate cinematicamente da gesti più lunghi e più ampi.

Per quanto riguarda la velocità massima, il confronto è significativo solo tra i gradi estremi di prominenzza. L'analisi della varianza è significativa sia per il locutore SG ($F(2,15)=10,05$ $p=0,0017$) che per il locutore CA ($F(2,10)=15,346$ $p=0,0009$). Ma il *test post hoc* indica per SG che il gesto di apertura delle atone non differisce in velocità da quello delle toniche, mentre è significativamente più lento rispetto alle focalizzate ($p=0,0005$), e che il gesto delle toniche è più lento di quello delle focalizzate ($p=0,006$). Per CA invece, il gesto di apertura delle atone è significativamente più lento sia di quello delle toniche ($p=0,0056$) sia di quello delle focalizzate ($p=0,0002$), mentre la differenza nella velocità massima di toniche e focalizzate è solo marginalmente significativa ($p=0,0852$). Le toniche dunque hanno una velocità intermedia, non significativamente diversa dalle atone nel soggetto SG, e diversa in modo solo marginalmente significativo rispetto alle focalizzate nel soggetto CA.

I due grafici nelle Figure 9 e 10 mostrano chiaramente l'esistenza di una covariazione positiva tra la massima ampiezza del gesto e la sua velocità massima: più un gesto è ampio, più è veloce.

Tale associazione è una caratteristica fondamentale dei sistemi oscillatori, ed è stato più volte osservato sia nei gesti articolatori propri del parlato (Kuehn e Moll, 1976) sia in altri tipi di movimenti (Cooke, 1980). I nostri dati indicano quindi che il gesto di apertura dalla consonante verso la vocale è governato da un sistema dinamico.

Un secondo fatto interessante che emerge dalla relazione di ampiezza e velocità riguarda le linee di regressione che interpolano separatamente i dati relativi ai diversi tipi di prominenzza sillabica. Sia per il locutore CA (Figura 9) che per il locutore SG (Figura 10), la pendenza della linea di regressione delle atone è più ripida. Una differenza nella pendenza della linea di regressione che correla ampiezza e velocità in sillabe con diverse tipologie accentuali era già stata notata da Ostry *et al.* (1983) e Kelso *et al.* (1985): nei loro dati, la pendenza della linea delle atone risultava essere la più ripida. Ovvero, nei gesti per le atone che, come abbiamo visto anche nei nostri risultati al paragrafo precedente, hanno minore ampiezza e velocità massima rispetto ai gesti con maggiori gradi di prominenzza, a un dato cambiamento in ampiezza corrisponde un maggior cambiamento in velocità rispetto alle toniche³.

³ È da notare che nel lavoro di Kelso e colleghi non vengono distinti tre gradi di accento ma solo due: sillabe *unstressed* vengono contrapposte a sillabe *stressed*; ma il contesto in cui le paro-

Figura 9. Soggetto CA. Velocità massima vs. ampiezza dell'Apertura Labiale. Velocità massima: atona <***{tonica, focalizzata}

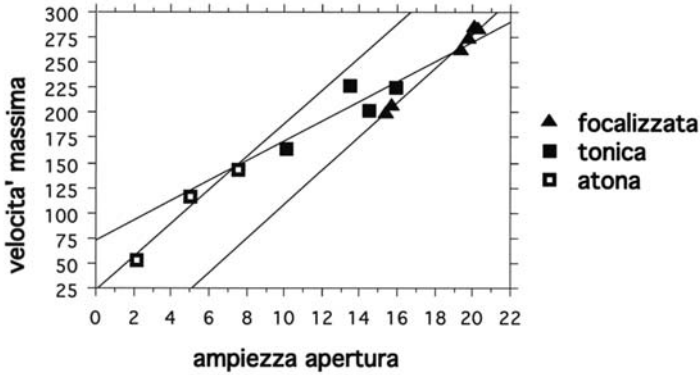
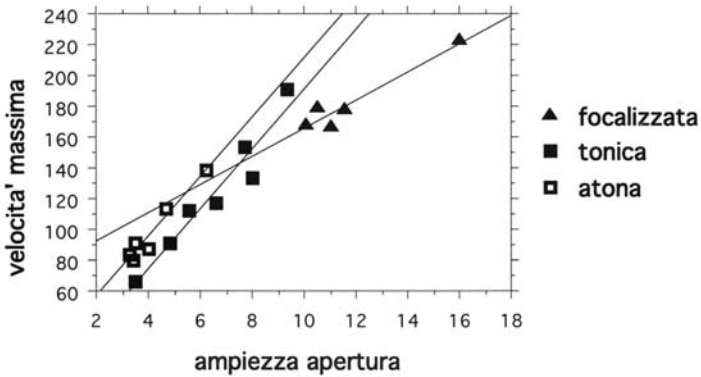


Figura 10. Soggetto SG. Velocità massima vs. ampiezza dell'Apertura Labiale (A T= atona, T= tonica, F= focalizzata) Velocità massima: {atona, tonica} <***focalizzata



Tale differenza, entro un sistema dinamico, è compatibile con un cambiamento nella rigidità del sistema, cioè nella sua frequenza naturale di oscillazione:

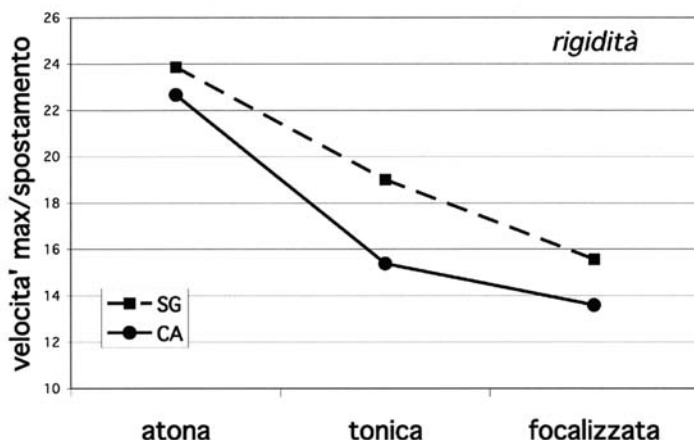
le accentate sono pronunciate lascia intendere che le sillabe *unstressed* siano portatrici anche di accento di frase, e dunque le sillabe categorizzate come *stressed* siano in realtà paragonabili alle sillabe che nel nostro esperimento sono focalizzate.

richiamandoci all'equazione dinamica presentata in (1), questa differenza è espressa dal parametro k . Da questa prospettiva, i nostri risultati confermano ancora una volta che siamo in presenza di un sistema dinamico, e suggeriscono che le sillabe atone sono prodotte da un sistema oscillatorio più rigido di quello che governa la produzione delle sillabe toniche/focalizzate.

3.2.1.3. Rigidità

La significatività di questo dato è confermata dall'analisi della varianza calcolata sul valore che esprime il rapporto tra la velocità massima e l'ampiezza, spesso usato come stima della rigidità. I risultati indicano un'alta significatività della differenza nel grado di rigidità tra i gesti di apertura di sillabe atone toniche e focalizzate: il gesto per le atone è quello che esibisce la rigidità maggiore per entrambi i soggetti (SG: $F(2,15)=44,18$ $p<0,0001$; CA: $F(2,10)=27,988$ $p<0,0001$). La rigidità diminuisce progressivamente dalle atone alle focalizzate per entrambi i locutori. Andando nel dettaglio, per il soggetto SG il *test post hoc* rivela che ogni confronto a coppia è altamente significativo: la rigidità delle atone è maggiore di quella delle toniche ($p<0,0001$) che è a sua volta maggiore di quella delle focalizzate ($p=0,0012$); per il soggetto CA, le atone sono significativamente diverse dalle toniche ($p=0,0003$) e dalle focalizzate ($p<0,0001$), mentre toniche e focalizzate non hanno rigidità significativamente diverse, anche se numericamente manifestano lo stesso andamento del soggetto SG (cfr Figura 11).

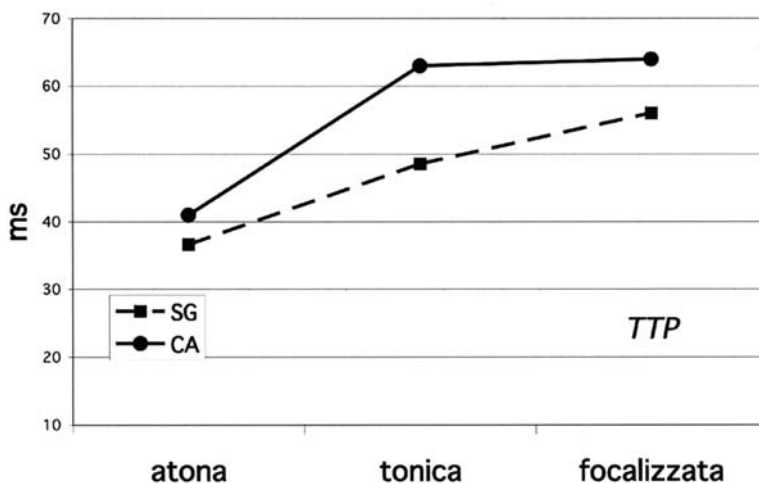
Figura 11. Valori di rigidità (velocità massima/ampiezza). Locutor e SG: atona >*** tonica >*** focalizzata. Locutore CA: atona >*** {tonica, focalizzata}



3.2.1.4. “Time-To-Peak velocity”

I valori di *TTP* (*Time-to-Peak velocity*) risultano speculari a quelli della rigidità (Figura 12). *TTP* esprime la durata del periodo di accelerazione, e indica l'intervallo temporale dall'*onset* del gesto al momento in cui il gesto raggiunge la sua massima velocità. Entro un sistema dinamico a massa-molla, la velocità massima viene raggiunta più tardi nei gesti che mostrano rigidità minore. I dati confermano che nelle sillabe focalizzate i valori di *TTP* sono maggiori di quelli delle atone e delle toniche. L'analisi della varianza indica una differenza altamente significativa in entrambi i soggetti (SG: $F(2,15)=44,18$ $p<0,0001$; CA: $F(2,10)=19,951$ $p=0,0003$). Il locutore SG mostra una progressione significativa dal valore di *TTP* che caratterizza il gesto della sillaba atona a quello della sillaba focalizzata (atona vs tonica: $p=0,01$; tonica vs focalizzata: $p=0,047$). Per il locutore CA, specularmente ai dati sulla rigidità, i valori sono progressivamente maggiori ma la significatività statistica è raggiunta solo nel confronto tra atona vs. tonica ($p=0,0004$) e atona vs. focalizzata ($p=0,0001$).

Figura 12. Durata dell'accelerazione (Time to Peak velocity). Locutore SG: atona <*** tonica <*** focalizzata. Locutore CA: atona <*** {tonica, focalizzata}



3.2.2. Struttura della sillaba: aperta vs. chiusa

Per verificare le proprietà cinematiche del gesto di apertura in relazione alla struttura sillabica, abbiamo analizzato i dati di sillabe CV e CVC con vocale [a] e

[i] prodotte dal locutore SG in una sola condizione accentuale, quella della massima prominente in *focus* contrastivo.

La durata acustica della sillaba aperta è significativamente diversa da quella della sillaba chiusa, sia per la sillaba con vocale [a] che per la sillaba con vocale [i] (rispettivamente: $F(1,11)=6,344$ $p=0,028$; $F(1,9)=11,688$ $p=0,0076$). Lo stesso vale per la durata della vocale: è significativamente più breve in sillaba chiusa che in sillaba aperta ([a]: $F(1,11)=5,736$ $p=0,035$; [i]: $F(1,9)=5,399$ $p=0,0452$). Si conferma quindi che esiste un accorciamento nella durata acustica della vocale in sillaba chiusa.

La minore durata acustica riflette una diminuzione nella durata articolatoria del gesto di apertura: questa è significativamente più breve nel gesto di apertura nelle sillabe chiuse, sia per la vocale [a] che per la vocale [i] ([a]: 108 vs 91 ms ($F(1,11)=5,373$ $p=0,04$); [i]: 98.3 vs 72 ms ($F(1,9)=5,744$ $p=0,04$).

Ma le altre variabili cinematiche considerate non sono significativamente diverse nelle due condizioni sillabiche, per nessuna delle due vocali: n è l'ampiezza del gesto, né la sua velocità massima, né la durata dell'accelerazione (*TTP*).

4. DISCUSSIONE – INTERPRETAZIONE IN TERMINI DI *TASK DYNAMICS*

In questa parte del lavoro cercheremo di rispondere alle seguenti domande:

- 1) l'insieme delle variazioni cinematiche indotte dall'accento e dalla struttura sillabica che abbiamo fin qui osservato sono compatibili con le predizioni del modello proposto dalla *Task Dynamics*? Se così fosse, tutte le differenze cinematiche osservate e le differenze di durata acustica dovrebbero conseguire dalla specificazione dei parametri dinamici di un oscillatore massa-molla.
- 2) Specificamente, in che modo possiamo derivare l'insieme delle variazioni cinematiche osservate dalla specificazione dei parametri dinamici espressi nell'equazione di moto (1)?
- 3) Se i dati cinematici sono compatibili con il modello della *Task Dynamics*, possiamo identificare un insieme di valori dei parametri dinamici che riporti in modo non ambiguo ad un solo meccanismo identificabile come strategia articolatoria alla base della produzione delle variazioni accentuali da un lato e delle variazioni indotte dalla struttura sillabica dall'altro?

Ricordiamo che le differenze di durata osservate superficialmente nel segnale acustico possono derivare entro un sistema dinamico da cambiamenti a) nel *target* (nello spostamento articolatorio soggiacente, il parametro x); b) nella rigidità del sistema (il parametro k); c) nel *target* e nella rigidità scalati proporzionalmente; d) nella coordinazione temporale (relazioni di fase) tra gesti contigui.

Cercheremo di fornire una risposta alle domande precedenti analizzando i dati cinematici e acustici separatamente per i singoli gradi della prominenzza e per la struttura sillabica.

4.1. *Atone vs. toniche*

Le vocali atone, rispetto alle toniche, sono acusticamente più brevi e il gesto vocalico di apertura che le caratterizza mostra le seguenti caratteristiche cinematiche relative alla variabile *AL* (Apertura Labiale): durata minore, ampiezza minore, velocità massima minore. Inoltre, il gesto mostra un *TTP* minore e una Rigidità maggiore. Queste proprietà dei gesti caratterizzano le sillabe atone e toniche di entrambi i soggetti⁴.

In un sistema dinamico, l'insieme di queste caratteristiche cinematiche è predetto da un cambiamento proporzionale nella specificazione del *target* e nella rigidità (cfr. § 2.4 e Figura 2c): ciò indica che le atone sono *globalmente* contratte rispetto alle toniche, e che le differenze di durata acustica osservate superficialmente sono la conseguenza diretta di tale contrazione.

4.2. *Toniche vs. focalizzate*

Per un confronto tra i due gradi più alti di prominenzza, bisogna distinguere le produzioni dei due soggetti. Per SG, il gesto di apertura delle sillabe toniche è più breve, meno ampio e meno veloce di quello delle sillabe focalizzate. Inoltre, mostra una rigidità maggiore e un *TTP* numericamente minore ma non significativamente diverso (48,57 vs 56 ms). La durata acustica è significativamente più breve.

Come per il confronto ai gradi minori di prominenzza, dunque, l'insieme delle caratteristiche cinematiche del gesto di apertura labiale in questo soggetto è predetto da un cambiamento proporzionale nella specificazione del *target* e nella rigidità: indica che le toniche sono contratte rispetto alle focalizzate, e che le differenze di durata acustica osservate superficialmente sono la conseguenza diretta di tale contrazione.

Per il secondo soggetto invece la differenza tra toniche e accentate è riportabile ad un diverso meccanismo soggiacente. Acusticamente le toniche sono più brevi delle focalizzate, e articolatoriamente si distinguono dalle focalizzate per avere: durata minore, ampiezza minore, velocità massima uguale⁵, *TTP* uguale e rigidità uguale.

⁴ La velocità massima del gesto di apertura per il locutore SG non è significativamente diversa nelle atone e nelle toniche. I valori assoluti però differiscono nella direzione predetta dal modello dinamico: la velocità delle atone è minore di quella delle toniche (98,84 vs. 123,62 mm/sec).

⁵ Ricordiamo che la velocità è significativa ad un livello più alto di probabilità: $p = 0,0852$.

Valori minori di durata e ampiezza del gesto a fronte di uguale *TTP*, velocità massima e rigidità suggeriscono un cambiamento nella sincronia di gesti contigui: nel caso delle sillabe focalizzate il gesto di *chiusura* da V a C viene attivato più tardi, e quindi permette al gesto di apertura di avere una realizzazione più piena, di attingere un *target* più ampio e di avere maggiore durata rispetto al gesto di apertura delle toniche. Questo spiega il fatto che il movimento di apertura abbia la stessa rigidità e la stessa velocità e impieghi lo stesso tempo per arrivare al punto di velocità massima nei due tipi di sillaba, ma che il movimento continui nel tempo e raggiunga un *target* maggiore solo nel caso delle sillabe focalizzate.

Per il soggetto SG, dunque, l'accento ai diversi livelli della gerarchia prosodica, è realizzato grazie a un aumento progressivo dei valori di tutte le variabili cinematiche e della variabile dinamica della rigidità. L'insieme di tali variazioni è predetto come conseguenza della diversa specificazione dei valori di uno stesso parametro dinamico: la *contrazione* (e il suo opposto, *l'espansione*); questo parametro è definito in termini di *Task Dynamics* come aumento proporzionale nei valori che specificano il *target* e la rigidità del gesto. La durata acustica progressivamente maggiore emerge dunque come conseguenza diretta della diversa specificazione di questo parametro dinamico in relazione ai diversi gradi di prominenza della sillaba (si veda la Figura 13, parte sinistra).

Nell'altro soggetto (CA) lo stesso meccanismo di espansione del gesto caratterizza solo il contrasto atona-tonica, ma non il contrasto tonica-focalizzata, che è realizzato mediante una posticipazione del gesto di chiusura (dunque per mezzo di una diversa coordinazione temporale tra gesti). A livello lessicale, l'opposizione tra sillabe atone e sillabe toniche è realizzata articolatoriamente con le stesse variazioni nei parametri cinematici e dinamici del soggetto SG: le toniche sono, rispetto alle atone, più lunghe, più ampie e più veloci e dinamicamente più rigide. E uguale è il meccanismo dinamico soggiacente che predice l'insieme delle variazioni parametriche osservate: la *contrazione*. A un livello più alto della gerarchia prosodica, l'opposizione tra sillabe toniche vs. focalizzate non è invece compatibile con lo stesso meccanismo, ma con meccanismo diverso che induce una realizzazione piena del gesto in termini di durata e di ampiezza senza modificarne la rigidità: una diversa sincronizzazione della fase di gesti contigui tale che l'attivazione del gesto di chiusura successivo (da V a C) 'slitta' rispetto alla fine del gesto di apertura e 'lascia più spazio', per così dire, alla piena realizzazione del gesto precedente (Figura 13, parte destra). Si può ipotizzare che l'adozione di due strategie diverse per realizzare gradi diversi del *continuum* di prominenza da atone a focalizzate – una espansione progressiva da una parte e uno slittamento nella sincronizzazione temporale di gesti contigui dall'altra – sia dovuto ai valori già alti di ampiezza e velocità massima del movimento raggiunti nella produzione delle sillabe toniche. I valori di ampiezza e velocità delle toniche di CA sono maggiori dei valori che SG mostra per le focalizzate (ampiezza toniche CA: 13,45 – ampiezza focalizzate SG: 11,84; velocità toniche CA: 205 – velocità focalizzate SG: 182,37). Il rapporto tra velocità e ampiezza è un

indice della rigidità del gesto, come abbiamo visto. Se la rigidità ha già raggiunto una soglia massima, per la velocità di eloquio a cui il soggetto ha parlato, allora la prominente maggiore del gesto per le focalizzate può essere realizzata con altri meccanismi dinamici: la posticipazione della fase del gesto di chiusura crea la possibilità di produrre un gesto di apertura più lungo, che raggiunge un *target* più elevato rispetto a quello delle toniche (ampiezza = 18,4 vs 13,4), ma con una differenza proporzionalmente inferiore rispetto a quella che distingue le toniche dalle atone (differenza di ampiezza: atona vs. tonica = 8,57; tonica vs. focalizzata = 4,96).

Figura 13. *Diversi meccanismi dinamici alla base della produzione di sillabe con gradi progressivamente crescenti di prominente: contrazione (a sinistra) ed espansione (a destra)*



4.3. *Sillabe aperte vs. sillabe chiuse*

I dati sui diversi tipi sillabici riguardano uno dei soggetti (SG), un grado di prominente (sillabe focalizzate) e le vocali [a] ed [i].

Il gesto di apertura che caratterizza le vocali nelle sillabe chiuse è più corto che nelle sillabe aperte, ma nessuna delle altre variabili cinematiche lo è: né l'ampiezza del gesto, né la sua velocità massima, né il *TTP*. In termini di *Task Dynamics* questi aspetti sono tutti predetti da un cambiamento nella sincronizzazione di gesti contigui, specificamente da un troncamento del gesto di apertura da parte del gesto seguente. Il gesto di chiusura, che caratterizza la consonante postvocalica, viene attivato anticipatamente rispetto alla fine del gesto di apertura: questo spostamento indietro nel tempo del gesto di chiusura lascia inalterate le caratteristiche di velocità, rigidità e velocità di accelerazione del gesto di apertura e determina solamente una diminuzione della sua durata. Il fatto che i dati cinematici non rivelino alcun cambiamento nell'ampiezza è compatibile con il fatto che il troncamento sia avvenuto nella parte di *plateau* del gesto di apertura (cfr. Figura 1c 'fase anticipata'), non influenzando in tal modo sui valori del *target* x_0 che risulta non significativamente diverso nei due tipi sillabici. Questi risultati trovano conferma anche in Zmarich *et al.*, 2006).

5. CONCLUSIONI

All'interno del quadro teorico della *AP* abbiamo applicato la nozione di gesto articolatorio alla descrizione di due diversi fenomeni prosodici: l'accento, visto nelle sue diverse categorie/gradazioni accentuali, e il cosiddetto 'accorciamento di compenso'. Entrambi questi fenomeni implicano variazioni nella durata acustica della vocale.

I nostri dati suggeriscono alcune considerazioni generali. Per prima cosa risulta confermata la validità del modello dinamico funzionale di Saltzman e colleghi qui esteso alla descrizione dinamica di fenomeni prosodici. La nozione di gesto articolatorio, inteso come unit à d'azione (fisica) e unit à di informazione (fonologica) ci consente di indagare i fondamenti dinamici di alcune variazioni di durata linguisticamente condizionate: accento e tipologia sillabica. Siamo partiti dall'idea che alle origini delle variazioni di durata acustica associate all'accento nelle sue varie gradazioni di prominente e alla diversa struttura sillabica vi siano processi di controllo articolatorio differenti. Basandoci anche su nostri dati precedenti, abbiamo ipotizzato che la minore durata acustica delle vocali atone rispetto alle vocali prominenti sia la conseguenza di un controllo diretto sul gesto articolatorio soggiacente: il gesto per le atone è globalmente diverso da quello per le sillabe prominenti, a partire dall'*onset* del gesto. La minore durata delle vocali in sillaba chiusa rispetto a quelle in sillaba aperta invece, secondo la nostra ipotesi, non è indotta da un controllo diretto che agisce sulla forma globale del gesto, ma è la conseguenza di una diversa coordinazione tra gesti contigui; specificamente è il risultato di un inizio anticipato del gesto di chiusura per la consonante tautosillabica che 'taglia' una parte del gesto di apertura della vocale precedente. L'indagine articolatoria e la nostra interpretazione dei dati cinematici in termini del modello dinamico confermano entrambe le ipotesi.

Per quanto riguarda l'accento, abbiamo visto che l'espansione globale e progressiva del gesto vocalico (da C a V) è la strategia unitaria che in un soggetto (SG) sottostà alla produzione di sillabe con grado crescente di accento. Questo fenomeno di espansione globale della vocale è coerente con l'idea tradizionale che "accenting reflects an overall increase in global articulatory effort" (Fowler, 1995). L'insieme di tali variazioni è predetto come conseguenza della diversa specificazione dei valori di uno stesso parametro dinamico: la *contrazione* (e il suo opposto, *l'espansione*); questo parametro è definito in termini di *Task Dynamics* come aumento proporzionale nei valori che specificano il *target* e la rigidità del gesto. La durata acustica progressivamente maggiore emerge dunque come conseguenza diretta della specificazione di questo parametro dinamico. Nell'altro soggetto (CA) lo stesso meccanismo di espansione del gesto caratterizza solo il contrasto atona-tonica, ma non il contrasto tonica-accentata, che è realizzato mediante una posticipazione del gesto di chiusura (dunque per mezzo di una diversa coordinazione temporale tra gesti).

In generale, le sillabe caratterizzate prosodicamente da tre livelli di prominente (atone, accentate lessicalmente, accentate intonativamente) hanno correlati

acustici e cinematici sistematicamente diversi. Quando una sillaba con accento lessicale assume un ulteriore livello di accento perché in *focus*, questa prominenza ulteriore che essa riceve è indotta pariteticamente sia dalle sue proprietà acustiche (allungamento) e cinematiche (ampiezza, velocità, durata) – che sono determinate intrinsecamente dalla specificazione di pochi parametri dinamici – sia dalla associazione con una specifica configurazione intonativa (*pitch accent*). Contrariamente a quanto avviene in inglese (Beckman e Edwards, 1994), le proprietà cinematiche della sillaba focalizzata non svolgono un ruolo ancillare nel segnalare la prominenza rispetto alle sue proprietà intonative.

Per quanto riguarda l' accorciamento di compenso, i nostri risultati sono in sintonia con altri studi precedenti sull'inglese (Munhall *et al.*, 1992) e sull'italiano (Vayra *et al.*, 1999; Zmarich *et al.*, 2006); confermano che l' accorciamento della durata vocalica nelle sillabe chiuse riflette una maggiore sovrapposizione tra i gesti di apertura e di chiusura per la consonante successiva in sillabe CVC rispetto alle sillabe CV, all'interno della stessa tipologia accentuale. L'attacco del gesto di chiusura è anticipato nelle sillabe chiuse. Entro l' ipotesi della coproduzione Vocale-Consonante, questo è un caso evidente di *blending* di due gesti orali che in un certo intervallo di tempo competono per l'uso di un articolatore in comune (la mandibola). L'esito del processo è il troncamento del gesto di apertura orale con conseguente effetto di riduzione acustica della durata vocalica.

BIBLIOGRAFIA

- Albano Leoni F., Cutugno F., Savy R., 1995, The Vowel System of Italian Connected Speech, in K. Elenius, P. Branderud (a cura di), *Proceedings of the XIIIth International Conference of Phonetic Sciences (ICPhS)*, Stockholm, 4, 396-399.
- Beckman M., 1986, *Stress and Non-stress Accent*, Dordrecht, Foris Publications.
- Beckman M., Edwards J., Fletcher M., 1992, Prosodic Structure and Tempo in a Sonority Model of Articulatory Dynamics, in G. Docherty, R. Ladd (a cura di), *Papers in Laboratory Phonology II. Segment, gesture and tone*, Cambridge U.K., Cambridge University Press: 68-86.
- Beckman M., Edwards J., 1994, Articulatory Evidence for Differentiating Stress Categories, in P. A. Keating (a cura di), *Papers in Laboratory Phonology III: Phonological Structure and Phonetic Form*, Cambridge U.K., Cambridge University Press: 7-33.
- Bertinetto P. M., 1981, *Strutture prosodiche dell'italiano*, Firenze, Accademia della Crusca.
- Browman C. P., Goldstein L. M., 1986, Toward an Articulatory Phonology, in "Phonology Yearbook", 3: 219-252.
- Browman C., Goldstein L., 1989, Articulatory Gestures as Phonological Units, in "Phonology", Volume 6: 201-251.
- Browman C., Goldstein L., 1990, Tiers in Articulatory Phonology, with Some Implications for Casual Speech, in J. Kingston, M. Beckman (a cura di), *Papers in Laboratory Phonology I: Between the Grammar and the Physics of Speech*, Cambridge U.K., Cambridge University Press: 341-376.
- Browman C. P., Goldstein L., 1992, Articulatory Phonology: An Overview, in "Phonetica", 49: 155-180.
- Browman C. P., Goldstein L., 1995, Dynamics and Articulatory Phonology, in R. Port, T. Van Gelder (a cura di), *Mind in Motion: Explorations in the Dynamics of Cognition*, Cambridge MA, MIT Press: 175-193.
- Browman C.P., Goldstein L., 2000, Competing Constraints on Inter-gestural Coordination and Self-organization of Phonological Structures, in "Bulletin de la Communication Parlée", 5: 25-34.
- Byrd D., Saltzman E., 2003, The elastic phrase: modeling the dynamics of boundary-adjacent lengthening, in "Journal of Phonetics", 31: 149-180.
- Byrd D., Kaun A., Narayanan S., Saltzman E., 2000, Phrasal Signatures in Articulation, in M. B. Broe, J. B. Pierrehumbert (a cura di), *Papers in Laboratory Phonology V*, Cambridge, Cambridge University Press: 70-87.
- Clements G. N., 1992, Phonological Primes: Features or Gestures?, in "Phonetica", 49, 3-4: 181-193.
- Chen M., 1970, Vowel Length Variation as a Function of the Voicing of the Consonant Environment, in "Phonetica", 22: 129-159.
- Cho T., 2006, Manifestation of Prosodic Structure in Articulation: Evidence from Lip Kinematics in English, in L. Goldstein, D. Whalen, C. Best (a cura di), *Laboratory Phonology 8*, Berlin/New York, Mouton de Gruyter: 519-548.

- Cooke J.D., 1980, The Organization of Simple Skilled Movements, in G. Stelmach, J. Requin (a cura di), *Tutorials in Motor Behavior*, Amsterdam, North Holland: 199-212.
- Dauer R., 1983, Stress-timing and Syllable-timing Reanalysed, in “Journal of Phonetics”, 11: 51-62.
- Farnetani E., 2003, The Supralaryngeal Articulation of Prominence in Italian Vowels, in P. Cosi, E. Magno Caldognetto, A. Zamboni (a cura di), *Voce canto parlato. Studi in onore di Franco Ferrero*, Padova, Unipress: 149-155.
- Farnetani E., Faber A., 1992, Tongue-jaw Coordination in Vowel Production: Isolated Words vs. Connected Speech, in “Speech Communication”, 11: 410-410.
- Ferrigno G., Pedotti A., 1985, ELITE: a Digital Dedicated Hardware System for Movement Analysis via Real-time TV Signal Processing, in “IEEE Trans. Biomed. Eng.”, BME 32: 943-950.
- Ferrigno G., Bettini F., Magno-Caldognetto E., 1999, Tecniche optoelettroniche per l’analisi dei disturbi articolatori, in A. Tronconi (a cura di), *Atti del 6° Convegno Nazionale Informatica, Didattica e Disabilità*, Andria (Bari), 4-6 novembre 1999: 84-88.
- Flemming E., 2001, Scalar and Categorical Phenomena in a Unified Model of Phonetics and Phonology, in “Phonology”, 18: 7-46.
- Fowler C., 1980, Coarticulation and Theories of Extrinsic Timing, in “Journal of Phonetics”, 8:113-133.
- Fowler C., 1981, A Relation between Coarticulation and Compensatory Shortening, in “Phonetica”, 38: 35-50.
- Fowler C., 1983, Converging Sources of Evidence on Spoken and Perceived Rhythms of Speech: Cyclic production of vowels in sequences of monosyllabic stress feet, in “Journal of Experimental Psychology: General”, 112: 386-412.
- Fowler, C., 1995, Acoustic and Kinematic Correlates of Contrastive Stress Accent in Spoken English, in F. Bell-Berti, L. Raphael (a cura di), *Producing speech: contemporary issues*, New York, AIP Press: 355-373.
- Gussenhoven C., 2004, *The Phonology of Tone and Intonation*, Cambridge, Cambridge University Press.
- Gay T., 1981, Mechanisms in the Control of Speech Rate, in “Phonetica”, 38: 148-158.
- Goldstein L., Fowler C. A., 2003, Articulatory Phonology: A phonology for Public Language Use, in N. O. Schiller, A. Meyer (a cura di), *Phonetics and Phonology in Language Comprehension and Production: Differences and Similarities*, Berlino, Mouton de Gruyter: 159-207.
- Halle M., 1983, On Distinctive Features and their Articulatory Implementation, in “Natural Language and Linguistic Theory”, 1: 91-105.
- Harrington J., Fletcher J., Roberts C., 1995, Coarticulation and the Accented/unaccented Distinction: Evidence from Jaw Movement, in “Journal of Phonetics”, 23: 305-322.
- Hawkins S., 1992, An Introduction to Task Dynamics, in G.J. Docherty, D.R. Ladd (a cura di), *Papers in Laboratory Phonology II. Gesture, Segment, Prosody*, Cambridge U.K., Cambridge University Press: 9-25.
- Keating P., 1990, The Window Model of Coarticulation: Articulatory Evidence, in J. Kingston, M. Beckman (a cura di), *Papers in Laboratory Phonology I: Between the*

- Grammar and the Physics of Speech*, Cambridge U.K., Cambridge University Press: 451-475.
- Kelso J.A.S., Tuller B, Harris K., 1983, A “Dynamic Pattern” Perspective on the Control and Coordination of Movement, in P. Mac Neilage (a cura di), *The production of speech*, New York, Springer-Verlag: 137-173.
- Kelso J.A.S., Vatikiotis-Bateson E., Saltzman E., Kay B., 1985, A Qualitative Dynamic Analysis of Reiterant Speech Production: Phase Portraits, Kinematics, and Dynamic Modeling, in “Journal of the Acoustical Society of America”, 77: 266-280.
- Krull D., 1989, Second Formant Locus Patterns and Consonant-vowel Coarticulation in Spontaneous speech, in “PERILUS (Phonetic Experimental Research at the Institute of Linguistics)”, University of Stockholm, 10: 87-108.
- Kuehn D., Moll K., 1976, A Cineradiographic Study of VC and CV Articulatory Velocities, in “Journal of Phonetics”, 4: 303-320.
- Ladefoged P., 1967, *Three Areas of Experimental Phonetics*, London, Oxford University Press.
- Lehiste I., 1970, *Suprasegmentals*, Cambridge MA, MIT Press.
- Lindblom B., 1963, *On Vowel Reduction*, The Royal Institute of Technology, Speech Transmission Laboratory, Stockholm, Sweden. Report No 29.
- Lindblom B., 1990, Explaining Phonetic Variation: A Sketch of the H e H Theory, in W. J. Hardcastle, A. Marchal (a cura di), *Speech Production and Speech Modelling*, Dordrecht, Kluwer: 403-439.
- Lindblom B., Agwele A., Sussman H. M., Cortes E. E., 2007, The effect of emphatic stress on consonant vowel coarticulation, in “Journal of the Acoustical Society of America”, 121 (6): 3802-3813.
- Lindblom B., Rapp K., 1973, Some Temporal Regularities of Spoken Swedish, in “Papers in Linguistics from the University of Stockholm”, 21: 1-59.
- Lindblom B. Lyberg B., Holmgren K., 1981, *Durational Patterns of Swedish Phonology. Do they Reflect Short-term Memory Processes?*, Bloomington Indiana, Indiana University Linguistics Club.
- Maddieson I., 1985, Phonetic Cues to Syllabification, in V. Fromkin (a cura di), *Phonetic Linguistics. Essays in Honor of Peter Ladefoged*, Orlando, Academic Press: 203-221.
- Munhall K., Fowler C., Hawkins S., Saltzman E., 1992, Compensatory Shortening in Monosyllables of Spoken English, in “Journal of Phonetics”, 20: 225-239.
- Magno Caldognetto E., Vaggés K., Zmarich C., 1995, Visible Articulatory Characteristics of the Italian Stressed and Unstressed Vowels, in *Proceedings of the XIII International Congress of Phonetic Sciences*, Stockholm, 1: 366-369.
- Nespor M., Vogel I., 2007, *Prosodic Phonology*, Berlin, Mouton de Gruyter.
- Öhman S.E.G., 1967, Numerical Model of Coarticulation, in “Journal of the Acoustical Society of America”, 41: 310-320.
- Oller D. K., 1973, The Effect of Position in Utterance on Speech Segments Duration in English, in “Journal of the Acoustical Society of America”, 54(5): 1235-1247.
- Ostry D., Keller E., Parush A., 1983, Similarities in the Control of the Speech Articulators and the Limbs: Kinematics of Tongue Dorsum Movement in Speech, in “Journal of Experimental Psychology: Human Perception and Performance”, 9: 637-651.

- Pierrehumbert J., 1990, Phonological and Phonetic Representation, in “Journal of Phonetics”, 18: 375-394.
- Perrier P., Lœvenbruck H., Payan Y., 1996, Control of the Tongue Movements in speech: The Equilibrium Point Hypothesis Perspective, in “Journal of Phonetics”, 24: 53-75.
- Saltzman E., 1986, Task Dynamic Coordination of the Speech Articulators: A Preliminary Model, in “Experimental Brain Research Series”, 15: 129-144.
- Saltzman E., 1995, Dynamics and Coordinate Systems in Skilled Sensorimotor Activity, in T. van Gelder e R. Port (a cura di), *Mind as Motion: Explorations in the Dynamics of Cognition*, Cambridge MA, MIT Press: 150-173.
- Saltzman E., Kelso J. A. S., 1987, Skilled Actions: A Task-Dynamic Approach, in “Psychological Review”, 94, 1: 84-106.
- Saltzman E., Munhall K. G., 1989, A Dynamical Approach to Gestural Patterning in Speech Production, in “Ecological Psychology”, 1, (4): 333-382.
- Saltzman E., Byrd D., 2000, Task-dynamics of Gestural Timing: Phase Windows and Multifrequency Rhythms, in “Human Movement Science”, 19: 999-526.
- Savy R., Clemente G., Lo Prejato M., 2005, Per una caratterizzazione e misura della riduzione vocalica in italiano, in P. Cosi (a cura di), *Misura dei parametri, Atti del 1° Convegno Nazionale dell'Associazione Italiana di Scienze della Voce, Padova, 2-4 dicembre, 2004*, Brescia, EDK Editore: 135-160.
- Summers W.V., 1987, Effects of Stress and Final-consonant Voicing on Vowel Production: Articulatory and Acoustic Analyses, in “Journal of the Acoustical Society of America”, 82: 847-863.
- Tisato G., Cosi P., Drioli C., Tesser F., 2005, Interface: Strumenti interattivi per l'animazione delle teste parlanti, in P. Cosi (a cura di), *Misura dei parametri, Atti del 1° Convegno Nazionale dell'Associazione Italiana di Scienze della Voce, Padova, 2-4 dicembre, 2004*, Brescia, EDK Editore: 817-846.
- Turvey M. T., 1977, Preliminaries to a Theory of Action with Reference to Vision, in R. Shaw, J. Bransford (a cura di), *Perceiving, acting and knowing: Toward an ecological psychology*, Hillsdale NJ, Lawrence Erlbaum Ass.: 211-265.
- Vatikiotis-Bateson E., 1987, *Linguistic Structure and Articulatory Dynamics: a Cross-language Study*, Tesi di dottorato, Indiana University.
- Vayra M., Fowler C., 1987, The Word Level Interplay of Stress, Coarticulation, Vowel Height and Vowel Position in Italian, in *Proceedings of the XI International Congress of Phonetic Sciences*, Tallinn, Estonia, URSS, Academy of Sciences of the Estonian S.S.R., IV: 24-27.
- Vayra M., 1997, Regole formali vs. regolarità fonetiche: Un caso di accorciamento compensativo in italiano, in F. Lo Piparo (a cura di), *Linguaggio e Cognizione, Atti del XXVIII Congresso della SLI, Palermo, 27-29 ottobre 1994*, Roma, Bulzoni: 357-388.
- Vayra M., Fowler C., 1992, Declination of Supralaryngeal Gestures in Spoken Italian, in “Phonetica”, 49, 1:48-60.
- Vayra M., Avesani C., Fowler C., 1999, On the Phonetic Bases of Vowel-consonant Coordination in Italian: a Study of Stress and “Compensatory Shortening”, in *Proceedings of the XIV International Congress of Phonetic Sciences*, S.Francisco, U.S.A., 1: 495-498.

- Vayra M., 2003, La coarticolazione nella Articulatory Phonology, in G. Marotta e N. Nocchi (a cura di), *La coarticolazione, Atti delle XIII Giornate di Studio del Gruppo di Fonetica Sperimentale, Pisa 28-30 novembre 2002*, Pisa, Edizioni ETS: 45-58.
- Zmarich C., Avesani C., Marchiori M., 2006, Coarticolazione e Accento, in V. Giordani, V. Bruseghini, P. Cosi (a cura di), *Atti del III Convegno Nazionale dell'Associazione Italiana di Scienze della Voce (AISV)*, Trento, 29-30/11-1/12/2006, Torriana (RN), EDK Editore srl: 2006, 5-15.
- Zmarich C., Gili-Fivela B., Perrier P., Savariaux C., Tisato G., 2006, Consonanti scempie e geminate in Italiano: studio acustico e cinematico dell'articolazione linguale e bilabiale, in V. Giordani, V. Bruseghini, P. Cosi (a cura di), *Atti del III Convegno Nazionale dell'Associazione Italiana di Scienze della Voce (AISV)*, Trento, 29-30/11-1/12/2006, Torriana (RN), EDK Editore srl: 151-163.

SITOGRAFIA

PRAAT: www.fon.hum.uva.nl/praat

