

L'ORGANIZZAZIONE TEMPORALE DEI GESTI VOCALICI E CONSONANTICI NELLE CONSONANTI SCEMPIE E GEMINATE DELL'ITALIANO

*Claudio Zmarich, °Barbara Gili Fivela, ^Pascal Perrier,

^Christophe Savariaux, *Graziano Tisato

*ISTC-CNR, sede di Padova, via Martiri della Libertà, 35137, Padova (I),

°Università del Salento, via Taranto 35, 73100 Lecce (I),

^ICP, 46, av. F. Viallet, 38031 Grenoble Cedex 1 (F)

zmarich@pd.istc.cnr.it, barbara.gili@ateneo.unile.it, perrier@icp.inpg.fr, savariaux@icp.inpg.fr,
tisato@pd.istc.cnr.it

1. SOMMARIO

Le differenze nella durata acustica e percettiva dei fonemi consonantici possono trovare spiegazioni in fatti articolatori di tipo differente, quali le diverse posizioni spaziali dei target e quindi delle ampiezze di movimento, le differenze relative alla velocità dei movimenti e la diversa sincronizzazione di gesti relativi a sequenze di fonemi.

Un caso particolare e molto interessante è rappresentato dalle consonanti geminate, cioè quei segmenti caratterizzati da durate fonologicamente lunghe. Questi segmenti rivestono particolare interesse perché possono essere confrontati con le loro controparti scempie per controllare varie ipotesi, da quelle che concernono la diversa posizione dei target articolatori a quelle relative a possibili cambiamenti nell'organizzazione temporale dei gesti articolatori in funzione della lunghezza fonologica dei segmenti da allocare nella catena fonica. Quest'ultima ipotesi si connette anche alla questione più generale riguardante il tipo di organizzazione prosodica propria di ciascuna lingua, in particolare per ciò che riguarda le sue caratteristiche ritmiche e la tipologia delle sue unità prosodiche. Per questi motivi, alcuni studiosi si sono interessati recentemente a quelle lingue che evidenziano un contrasto fonologico di lunghezza nel loro sistema consonantico, come il giapponese (Smith, 1995; Löfqvist, 2005; Löfqvist, 2006), lo svedese (Löfqvist, 2006) o l'italiano (Smith, 1995; Zmarich & Gili Fivela 2005, Gili Fivela et alii, 2007, Zmarich et alii, 2007), e che dal punto di vista genealogico e tipologico sono molto diverse tra loro.

Nel sistema fonologico dell'italiano si fa un uso pervasivo del contrasto di lunghezza per le consonanti omorganiche in posizione intervocalica interna alla parola. Studi acustici e percettivi hanno dimostrato da molto tempo che la durata delle geminate è all'incirca doppia della durata delle scempie e che la durata della vocale precedente contribuisce a distinguere scempie e geminate (Bertinetto, 1981). Studi recenti di tipo articolatorio sulla cinematica delle labbra e della lingua nella produzione di consonanti scempie, geminate e in gruppo consonantico raccolti con il sistema EMA presso l'ICP di Grenoble (F) su un'ampia tipologia di gesti labiali e linguali hanno riscontrato [Gili Fivela et alii, 2007] che il gesto consonantico realizzato da due parlanti italiani (Nord-Ovest e Nord-Est) non sembra sovrapposto (*overlapped*) al gesto vocalico contestuale in "mim(m)a", come predetto dal modello coarticolatorio "V-to-V" di Öhman (1967) e riscontrato da Smith (1995). In realtà l'organizzazione temporale reciproca di eventi articolatori quali l'istante di massima costrizione bilabiale rispetto all'inizio e alla fine del movimento della lingua per la vocale e l'intervallo temporale tra i due target vocalici ha richiesto l'ibridazione del modello di Öhman con il modello di organizzazione temporale "V+C" di Browman & Goldstein (1986).

Per quanto riguarda l'ipotesi dell'esistenza di target articolatori più o meno "distanti" per scempie e geminate (Löfqvist & Gracco, 1997; Löfqvist, 2005), i risultati sembrano puntare verso l'esistenza di target articolatori differenziati per posizione in funzione di combinazioni particolari di articolatori e tipi di gesti (di apertura o chiusura). Il dato più forte a favore dell'ipotesi riguarda la velocità significativamente maggiore, con il picco più vicino al rilascio, per il gesto di apertura delle geminate alveolari in entrambi i soggetti.

2. INTRODUZIONE

Nel sistema fonologico dell'italiano la maggioranza delle consonanti presenta un contrasto di lunghezza in posizione intervocalica interna alla parola. Le caratteristiche fonetiche delle consonanti geminate dell'italiano sono state presentate in una serie di studi, specialmente per quanto riguarda le caratteristiche acustiche e uditive. Per esempio, molti studi hanno stabilito che la durata delle geminate è all'incirca doppia della durata delle corrispondenti scempie e che la durata della vocale che precede immediatamente la consonante geminata è più breve della vocale che precede la consonante scempia (Bertinetto, 1981; Esposito & Di Benedetto, 1999).

Il primo studio di tipo cinematico sulle consonanti geminate dell'italiano era focalizzato sull'organizzazione temporale dei gesti articolatori consonantici rispetto ai gesti articolatori delle vocali adiacenti (Smith, 1995). L'autrice concludeva la sua analisi affermando che la strategia articolatoria sottostante la produzione delle geminate in italiano era meglio rappresentata dal modello *Vowel-to-Vowel*, proposto da Öhman (1967) invece che dal modello *Vowel & Consonants* di Browman and Goldstein (1986).

Altri studi di tipo cinematico considerano le geminate sotto un altro punto di vista, chiedendosi se siano meglio interpretate da una rappresentazione monosillabica o da una eterosillabica (Gili Fivela & Zmarich, 2005; Zmarich & Gili Fivela, 2005). Dopo aver analizzato le proprietà spaziali e temporali dei movimenti del labbro inferiore nella produzione di consonanti scempie, geminate e in gruppo consonantico (intervocalico), gli autori trovarono che il gesto di apertura (dalla chiusura consonantica alla vocale successiva) aveva durata, ampiezza, profilo di velocità (cfr. *Time-to-Peak*) simili al gesto di apertura del nesso consonantico, ed entrambi erano maggiori rispetto al gesto di apertura della consonante scempia. Inoltre la rigidità del gesto (cfr. *stiffness*) era uguale a quella del nesso consonantico e inferiore a quella per la consonante scempia. Sulla base di tali evidenze, gli autori concludevano per una rappresentazione eterosillabica delle consonanti geminate, grazie alla loro somiglianza articolatoria con i gruppi consonantici (sicuramente eterosillabici).

Questo lavoro prende le mosse dalle analisi comparative menzionate precedentemente tra le consonanti scempie e quelle geminate dell'italiano. Lo scopo del presente studio è quello di discutere i dati tratti da un insieme più ampio di gesti articolatori (delle labbra e della lingua) e testare in questo modo direttamente le ipotesi avanzate finora sull'organizzazione articolatoria delle consonanti geminate. Il primo obiettivo è quello di testare direttamente l'ipotesi di Smith (1995) sull'organizzazione temporale dei gesti nelle coppie di consonanti scempie/geminate dell'italiano, il secondo obiettivo è quello di raccogliere dati che permettano di ricostruire la posizione "virtuale" del target articolatorio per le consonanti scempie e geminate, per scoprire se è vero che i parlanti italiani realizzano le consonanti geminate come se cercassero di raggiungere un target "virtuale" più alto dal punto di vista spaziale di quello delle consonanti scempie (Löfqvist, 2005).

2.1 L'organizzazione temporale dei gesti articolatori

Secondo Smith (1995), nelle lingue a isocronia sillabica come l'italiano, la durata dell'intervallo temporale tra i due nuclei di sillabe successive non dipende dal numero delle consonanti intermedie (VCV=VCCV, v. fig.1).

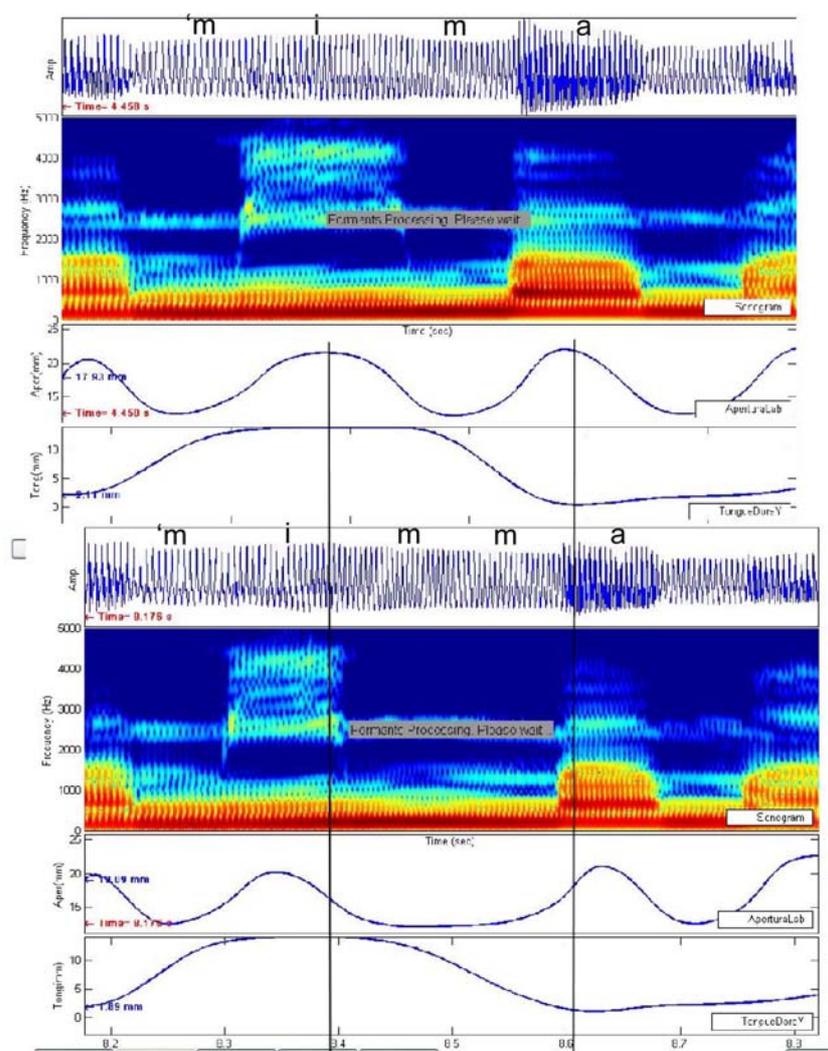


Fig. 1. Nella visualizzazione di *Interface*, sono mostrate due rappresentazioni dei segnali acustici (forma d'onda e sonogramma) e dei segnali cinematici (Apertura Labiale, dorso della lingua) relativi alle frasi "(richiam)a 'mima/mimma marca(tamente). I segnali cinematici relativi alla consonante scempia (sopra) e geminata (sotto) sono stati allineati sul target articolatorio della prima vocale (linea verticale sinistra).

Nel suo esperimento 3 parlanti italiani producevano consonanti scempie e geminate in un contesto V /i/-/a/ (per es. "mima" vs. "mimma"). Secondo Smith (1995), il *timing* di due vocali successive è controllato indipendentemente dalla consonante intervocalica (scempia

o geminata), e l'intervallo temporale tra il bersaglio cinematico di V1 e il bersaglio cinematico di C(C) non varia al variare della lunghezza della consonante. Ciò accade perché nella geminata il gesto di chiusura incomincia prima e il gesto di apertura finisce dopo (sovrapponendosi alle vocali contestuali), e il punto mediano del plateau di tenuta della chiusura consonantica è uguale a quello della scempia. Per Smith (1995) il modello che meglio spiega questo tipo di organizzazione articolatoria dell'italiano è il *V-to-V model* (Öhman, 1967), dove il gesto articolatorio per la consonante è considerato come sovrapposto al gesto da vocale a vocale, e quest'ultimo forma la base ritmica. Come esempio di applicazione di un modello alternativo, la Smith (1995) cita il giapponese, che non è una lingua a isocronia sillabica, e i cui dati sperimentali sono meglio spiegati dal modello *Vowel & Consonant* di Browman & Goldstein (1986). In effetti, nel giapponese il gesto vocalico (eseguito dal dorso della lingua) parte alla fine di quello consonantico, e di conseguenza la durata dell'intervallo temporale tra i due nuclei vocalici e la collocazione temporale del punto di massima costrizione consonantica dipendono entrambi dal numero di consonanti intermedie, come se ogni consonante aggiuntiva spostasse sempre un po' più a destra l'ultima vocale (vedi fig. 2; per l'organizzazione temporale del giapponese, cfr. anche Löfqvist, 2006).

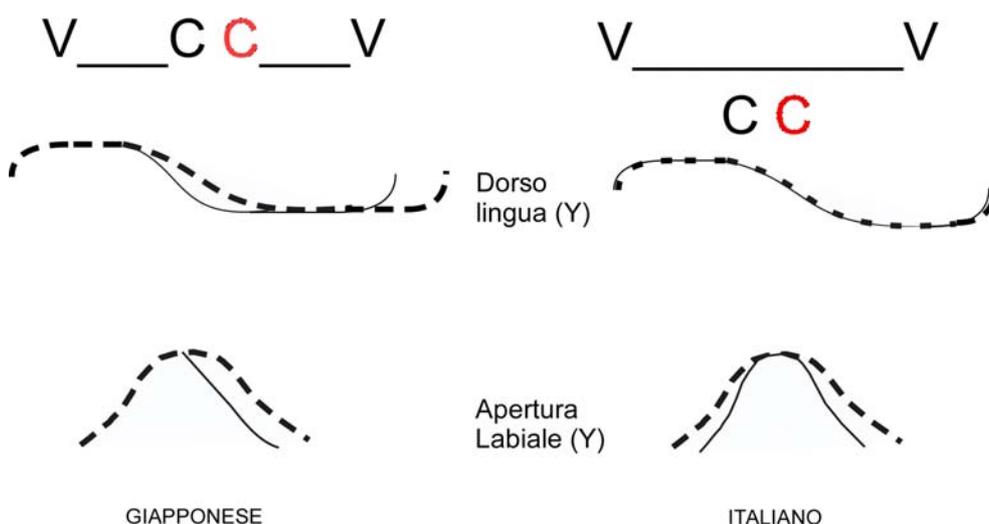


Fig. 2. I due possibili modelli di coordinazione temporale tra i gesti consonantici e vocalici secondo Smith (1995) per C. scempie e geminate: il modello *V & C* esemplificato dal giapponese (a sin.), e il modello *V-to-V* esemplificato dall'italiano (a destra).

L'ipotesi di Smith (1995) fu testata per l'italiano da Zmarich e Gili-Fivela (2005), che analizzarono le caratteristiche cinematiche di consonanti nasali realizzate in contesto vocalico /a/. I risultati furono abbastanza inconcludenti, nel senso che uno dei due soggetti sembrava aderire maggiormente al modello di Öhman (1967) e l'altro al modello di Browman e Goldstein (1986). Ciò potrebbe essere dovuto a strategie articolatorie idiosincratiche o all'influenza di un diverso sostrato linguistico dei due soggetti (Nord-Ovest e Nord-Est). Tuttavia la spiegazione più semplice fa appello a un artefatto metodologico, poiché le misure dell'apertura labiale li utilizzate, pur essendo correlate alla

produzione della vocale in termini maggiore apertura (cfr. Magno-Caldognetto, Vagges & Zmarich, 1995), non misurano direttamente il movimento del dorso della lingua (che anzi continua a muoversi anche quando la bocca è chiusa, cfr. Löfqvist, 2006), e quindi possono non essere sufficientemente precise per cogliere sottili differenze nell'organizzazione spazio-temporale dei gesti consonantici e vocalici.

2.2 La strategia articolatori nel raggiungimento del target consonantico

Per quanto riguarda la differenza nelle strategie articolatorie impiegate per la produzione delle consonanti scempie e geminate, Löfqvist (2005) propose che i parlanti possono controllare la durata del contatto occlusivo tra gli articolatori, e quindi ottenere quelle differenze di durata acustica relative all'intervallo di tenuta dell'occlusione che dal punto di vista percettivo veicolano maggiormente la distinzione di lunghezza fonologica, regolando la posizione spaziale del target articolatorio. Dato che la misurazione diretta del punto spaziale di contatto tra gli articolatori (per es. labbro inferiore e labbro superiore in un'occlusiva bilabiale) non rivela nessuna differenza significativa, cioè le labbra s'incontrano e restano unite più o meno alla stessa altezza per le consonanti scempie e geminate, per verificare l'ipotesi teorica bisogna supporre che i) il target possa essere virtuale, cioè collocato anche al di là della superficie dell'articolatore passivo (volta palatina o labbro superiore) e ii) possa essere misurato. Una di queste misure sfrutta la correlazione positiva relazione tra ampiezza e velocità massima in un movimento: *a parità di durata*, spostamenti maggiori vengono eseguiti a velocità più alte. Dunque secondo Löfqvist (2005), se si trovasse che nelle geminate i picchi di velocità dei gesti di chiusura e di apertura occorrono più vicini ai confini acustici e hanno valori più alti di quelli delle corrispondenti scempie, questo dimostrerebbe indirettamente l'esistenza di un target articolatorio delle geminate posto più in alto di quello delle scempie.

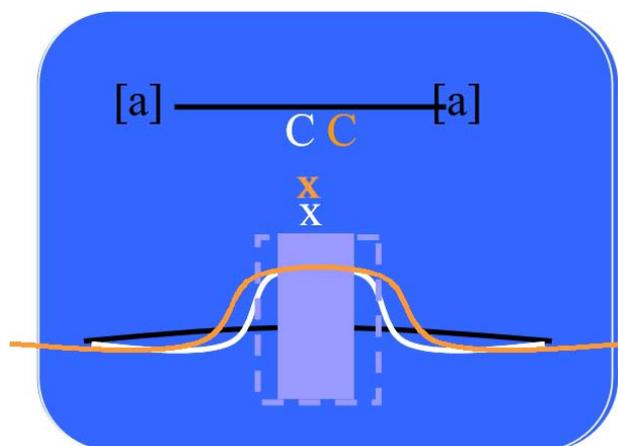


Fig. 3. Il gesto consonantico C(C) si sovrappone al gesto vocalico (a→a). Nella geminata (arancione) il target articolatorio (CC) è più alto che per la scempia (C) (bianco), e il periodo di contatto tra l'articolatore attivo e passivo è maggiore (maggiore durata della chiusura acustica, cfr. rettangolo tratteggiato grigio rispetto al pieno)

E' da notare che una strategia articolatoria che si basa solo sulla diversità dell'ampiezza del movimento per implementare variazioni di durata, presenta, a parità di altri fattori, il

vantaggio dell'economia esplicita (perchè non ricorre alla programmazione temporale esplicita).

I suoi dati sperimentali, tratti dalla cinematica del labbro inferiore di parlanti svedesi e giapponesi, non evidenziavano però differenze significative tra consonanti scempie e geminate nella velocità massima (a parte un soggetto svedese). Quindi Löfqvist (2005) concludeva che i soggetti non cambiano solo la posizione del punto finale del movimento, (il target), ma intervengono anche sull'organizzazione temporale, per es. rallentando il movimento allo scopo di mantenere più a lungo il contatto tra i due articolatori.

3. MATERIALI E METODI

3.1 Corpus

Il corpus utilizzato in questa indagine è formato da parole e pseudo-parole CVC(C)V generate con le seguenti regole: a) l'accento cade sulla prima sillaba; b) le due possibili consonanti (occlusive sonore, dentali e velari, e laterali dentali) si oppongono per la lunghezza (per es. /d, l/ vs /dd, ll/); c) il contesto vocalico è sempre la /a/. A queste regole se ne è aggiunta una quarta per la creazione di *cluster* consonantici in cui le consonanti scempie dentali, la liquida laterale e l'occlusiva, comparissero a turno in ciascuna delle due posizioni di un cluster (/dl/, /ld/) (questa parte è già stata oggetto dello studio di Zmarich *et alii* (2007) e non sarà ulteriormente approfondita in questa sede). Si voleva in questo modo permettere il confronto tra le caratteristiche di un segmento scempio, di una geminata e del primo o del secondo segmento di un *cluster* analizzato come eterosillabico, con lo scopo di acquisire dati relativamente allo status mono o bifonemico delle geminate. Per quanto riguarda infine la verifica dell'ipotesi di Smith (1995) sull'organizzazione temporale di consonanti e vocali, sono state impiegate parole con consonanti nasali bilabiali e vocali /i/ ed /a/, come ad esempio nella parola "mim(m)a". Le complessive 12 parole e pseudoparole compaiono in una frase cornice in cui il target è seguita da una parola che comincia con la stessa consonante (per es., 'richiama *lal(l)a* latamente'), e che serve ovviamente a minimizzare possibili influenze coarticolatorie dal contesto non sotto studio. La lista completa del corpus è la seguente: "màma, màmma, làla, làlla, (gàga, gágga, làlda, dàdla), mima, mimma, dàda, dàdda". In questa sede non verranno analizzate le parole con nessi consonantici o con le occlusive velari. Si può notare come tutte le strutture analizzate costituiscono parole attestate o parole possibili della lingua italiana.

3.2 Soggetti

Le frasi in esame sono state acquisite complessivamente da quattro soggetti. In questo lavoro si presentano i risultati relativi solo a due di loro, entrambe donne e laureate (età: 30-35 anni), una originaria della provincia di Padova (NE) e l'altra di Torino (NW), dichiaratamente esenti da qualsiasi tipo di disturbo o disordine di voce, parola e linguaggio, presente e passato. Le frasi sono state randomizzate e presentate a gruppi di 3 su grandi pannelli posizionati a una distanza di circa due metri dal soggetto. Ai soggetti è stato chiesto di leggere ogni frase prima a velocità normale e subito dopo alla massima velocità sostenibile senza perdere di precisione. Per ogni soggetto sono state ottenute 10 ripetizioni di ciascuna delle 12 frasi a velocità normale e altre 10 a velocità elevata, per un totale di 240 enunciati (120+120) ed un totale complessivo di 960.

3.3 Strumentazione e procedura sperimentale

Il segnale audio è stato acquisito su registratore DAT tramite un microfono professionale, mentre i dati cinematici sono stati ottenuti con il sistema "Electro Magnetic

Articulograph” o EMA AG200 (© Carstens), nella versione 2D in uso presso il laboratorio dell’ICP a Grenoble (F).

In figura 3 si può vedere la configurazione dei sensori utilizzati (e che sono attaccati con colla medica). I sensori alla base del naso e tra gli incisivi superiori servono come riferimento per compensare i movimenti della testa. Quelli posti alla base del confine tra i due incisivi inferiori costituiscono il riferimento per l’articolazione della mandibola. Gli altri sensori erano posizionati nei seguenti punti anatomici: due al centro del labbro inferiore e superiore, e quattro sulla superficie della lingua, a distanza di circa un centimetro l’uno dall’altro, partendo da quello più anteriore, posizionato un centimetro dopo l’apice, e proseguendo verso la radice.

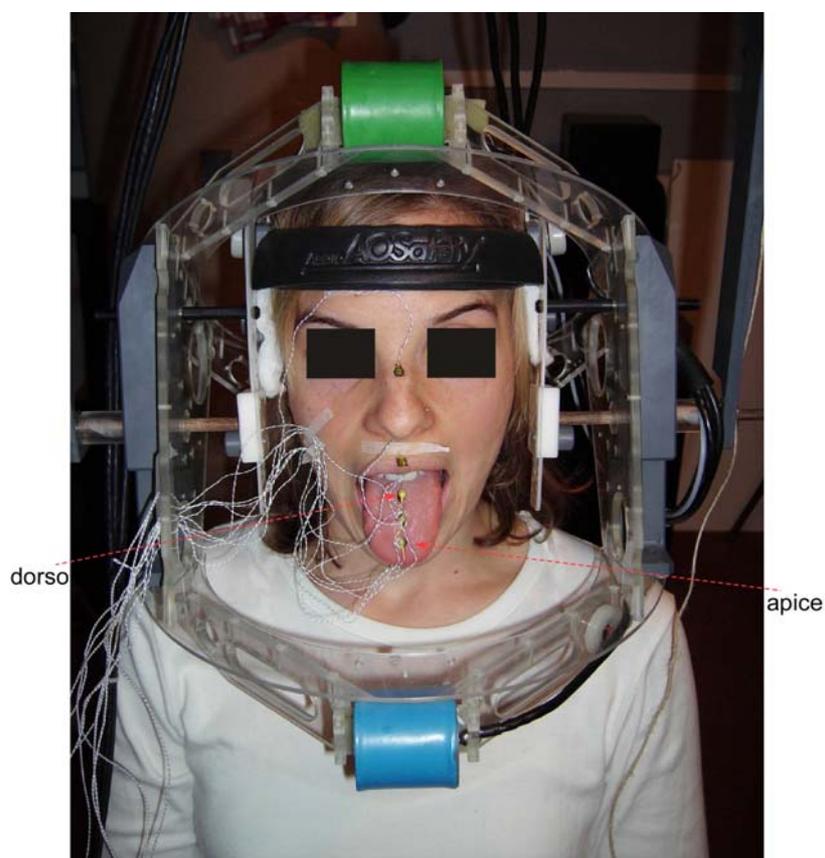


Figura 3: Posizione dei sensori durante la registrazione effettuata nei laboratori dell’ICP di Grenoble, mediante l’*Electro Magnetic Articulograph – 2D*, o EMA (© Carstens).

I dati discussi in questa sede riguardano le misurazioni relative al sensore incollato vicino all’apice della lingua, per le consonanti dentali, mentre per le consonanti labiali si sono utilizzati i sensori posti al centro del labbro superiore e inferiore, derivando la misura della distanza verticale, chiamata Apertura Labiale. Per le vocali, si è utilizzato il sensore incollato al dorso della lingua. Con una buona calibrazione e facendo attenzione che tutti i

sensori siano fra loro allineati lungo il sulcus centrale della lingua, si può ottenere una risoluzione spaziale del sistema con un margine di errore di 0.5 mm.

All'inizio di ogni seduta individuale di registrazione, un piccolo pannello rigido di forma rettangolare su cui erano stati attaccati due sensori è stato appoggiato sulla chiostr dentaria inferiore ed il soggetto è stato invitato a morderlo delicatamente per registrare la posizione della linea di occlusione. Alla fine della registrazione tutti i dati cinematici sperimentali sono stati corretti per il movimento della testa e ruotati fino a far coincidere la linea di occlusione con l'asse delle x . Questa rotazione è necessaria per ottenere un sistema di coordinate uniforme per tutti i soggetti.

3.4 Test uditivo

Mentre nell'italiano standard la durata delle consonanti è distintiva, diversi dialetti parlati nel Nord-Italia non fanno uso di geminate e quindi non sfruttano l'opposizione scempia/geminata. Per questo motivo, in via preliminare abbiamo deciso di verificare con un test di percezione se i nostri due soggetti avevano reso adeguatamente l'opposizione di lunghezza consonantica. Il test ha permesso di individuare le produzioni che risultavano chiaramente attribuibili all'una o all'altra classe di lunghezza. Le analisi acustiche e cinematiche sono state concentrate solo sulle consonanti riconosciute correttamente come scempie o geminate, per mettere in luce quei correlati acustici e articolatori associati al contrasto di lunghezza.

Una selezione dei 480 stimoli prodotti dai due parlanti è stata impiegata per il test di percezione. Per ciascun parlante, sono state scelte due ripetizioni prodotte a velocità normale e tutte le ripetizioni prodotte a velocità rapida (per un totale di 240 stimoli, nessuno dei quali con cluster consonantico). Un *panel* di cinque soggetti provenienti da Torino e cinque da Lecce e Taranto ha preso parte al test di percezione. Con l'ausilio del software *Perceval* (sviluppato presso il laboratorio LPL di Aix-en-Provence, Francia), i soggetti dovevano ascoltare i file audio che contenevano le parole target e giudicare se nelle parole erano presenti consonanti scempie o geminate. I soggetti ascoltavano ogni stimolo per due volte, e, dopo un segnale sonoro, avevano 5 secondi per rispondere, scegliendo tra due opzioni che venivano visualizzate sullo schermo (ad esempio, 1) 'lala' e 2) 'lalla'). Il test prevedeva una fase di addestramento nella quale ai soggetti erano presentati 5 stimoli scelti a caso tra quelli del corpus. Se gli stimoli erano riconosciuti in modo errato da almeno 2 soggetti di entrambe le varietà, o almeno 4 soggetti di una sola varietà, gli stessi stimoli non erano ulteriormente considerati per l'analisi acustica e cinematica.

3.5 Analisi acustica e cinematica

Dal punto di vista acustico, sono stati individuati i confini e gli intervalli corrispondenti ai segmenti che componevano le pseudo-parole target del corpus. Le analisi sono state effettuate grazie al software *PRAAT* (P.Boersma e D. Weenink, Institute of Phonetic Sciences of the University of Amsterdam). I risultati sono stati poi importati nell'ambiente di analisi cinematica. Le elaborazioni cinematiche sono state eseguite impiegando il software *Interface*, che è un software interattivo, creato da G. Tisato per l'elaborazione e l'analisi dei dati cinematici (cfr. Tisato, Cosi, Somnavilla, Zmarich, 2007). Delle due dimensioni spaziali in cui è ricostruito il movimento di labbra e lingua, quella verticale e quella longitudinale, è stato scelto come oggetto di analisi la dimensione verticale, più pertinente a rivelare possibili differenze per quanto riguarda le differenze nei movimenti delle labbra, del dorso e dell'apice della lingua, considerati separatamente nelle vocali e nelle consonanti occlusive e laterali che comparivano come scempie o geminate.

In figura 4 si può vedere una schermata tipica di *Interface*, che illustra l'andamento temporale di una serie di parametri acustici e cinematici. Dall'alto in basso si distinguono: il segnale acustico (sez. superiore), relativo a parte dell'enunciato contenente la parola target “(richiam)a 'lalla la(tamente)””; il sonogramma a banda larga; il segnale cinematico relativo all'ampiezza della componente verticale del movimento dell'apice della lingua; il segnale cinematico relativo alla velocità istantanea del movimento dell'apice della lingua.

Sulla curva relativa all'andamento dell'ampiezza verticale dell'apice della lingua abbiamo evidenziato il gesto di chiusura (dalla posizione aperta assunta per la prima /a/ di “lalla” all'occlusione relativa alla consonante intervocalica), e di apertura (dall'occlusione relativa alla consonante intervocalica alla posizione aperta relativa alla seconda /a/ di “lalla”), mentre sulla curva relativa alla velocità istantanea del movimento dell'apice della lingua abbiamo evidenziato (con un circolo pieno) i picchi di velocità del movimento di chiusura e di apertura. In modo analogo a quanto eseguito per es. da Smith (1995), l'inizio e la fine di ogni gesto sono stati collocati nei punti in cui la velocità dell'articolatore scendeva al di sotto del 10% della velocità massima raggiunta dallo stesso gesto (v. Fig.4).

Per ciascuno dei suddetti gesti abbiamo calcolato l'ampiezza (mm), e la durata (ms), che risultano automaticamente dalla differenza nella posizione verticale tra il punto iniziale e quello finale del movimento, ed abbiamo inoltre calcolato per ciascun gesto la velocità massima istantanea (mm/s).

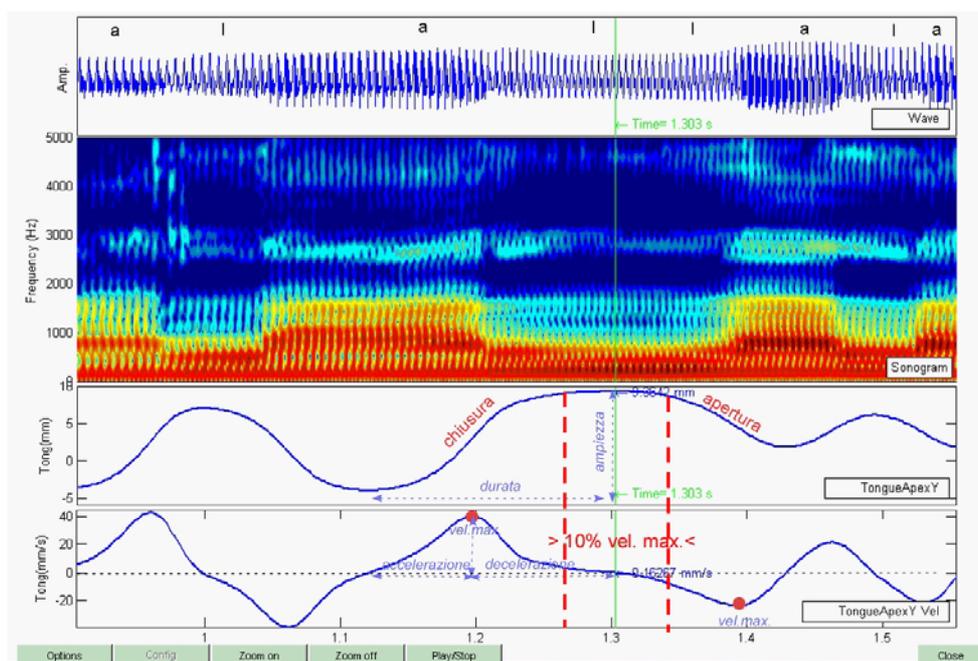


Figura 4. Schermata tipica di *Interface*, che riporta, dall'alto verso il basso, la forma d'onda del segnale acustico relativo alla frase “(richiam)a 'lalla la(tamente)””, il sonogramma a banda larga, il segnale cinematico relativo all'ampiezza della componente verticale del movimento dell'apice della lingua (mm); il segnale cinematico relativo alla velocità istantanea della componente verticale del movimento dell'apice della lingua (mm/s). Sulle curve cinematiche sono visibili i tipi di misurazioni effettuate.

In fig. 5 e fig. 6 vengono illustrati gli intervalli temporali calcolati per poter distinguere tra i due modelli alternativi di organizzazione temporale dei gesti articolatori:

1. tra il target articolatorio della prima vocale /a/ e il centro del ciclo consonantico /m(m)/. Laddove per la consonante la fine del gesto di chiusura non coincideva con l'inizio del gesto di apertura, il centro era collocato a metà del plateau di tenuta articolatoria, a bassa o nulla velocità. Se per l'italiano fosse valido il modello di Öhman (1967), questo intervallo nelle strutture VCCV dovrebbe essere uguale a quello per VCV (vedi Fig.5, n. 1).
2. tra il centro del ciclo consonantico /m(m)/ e il target articolatorio della seconda vocale /i/. Se per l'italiano fosse valido il modello di Öhman (1967), nelle strutture VCCV anche questo intervallo, che è speculare al precedente, dovrebbe essere uguale a quello per VCV (vedi Fig.5, n. 2).
3. tra il target articolatorio della prima vocale /i/ (cioè il punto più alto del dorso della lingua) e il target della seconda vocale /a/ (cioè il punto più basso del dorso della lingua): secondo il modello di Öhman (1967), nelle strutture VCCV e VCV dovrebbero essere uguali (vedi Fig.5, n. 3).
4. tra il target della prima vocale /i/ e il momento iniziale della chiusura per la consonante /m(m)/: secondo il modello di Öhman (1967), nelle strutture VCCV dovrebbe essere inferiore a quello per VCV (vedi Fig.6, n.4).
5. tra il momento finale del gesto di apertura della consonante /m(m)/ e la posizione del target per la seconda vocale (a): secondo il modello di Öhman (1967), nelle strutture VCCV dovrebbe essere inferiore a quello per VCV (vedi Fig. 6, n. 5).

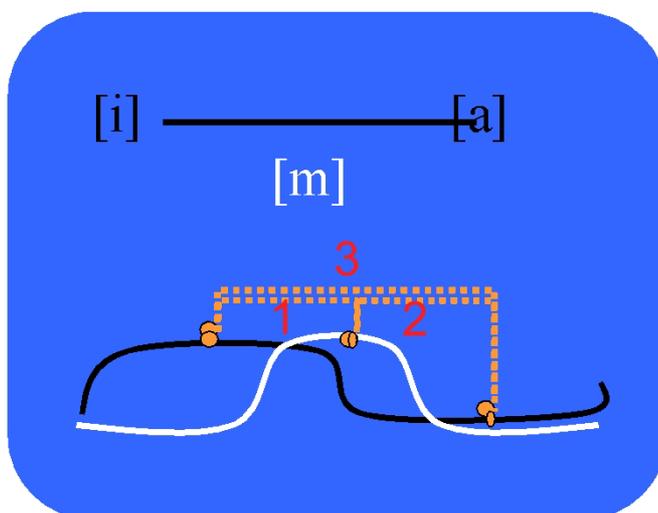


Fig 5. Schema delle misure effettuate per verificare il modello di coordinazione temporale C-V di Öhman (1967) applicato a “mim(m)a”. La linea nera si riferisce al movimento del dorso della lingua da [i] ad [a], mentre la linea bianca si riferisce al gesto dell’Apertura Labiale per la realizzazione della consonante [m]/[mm]. Vedi testo per la spiegazione.

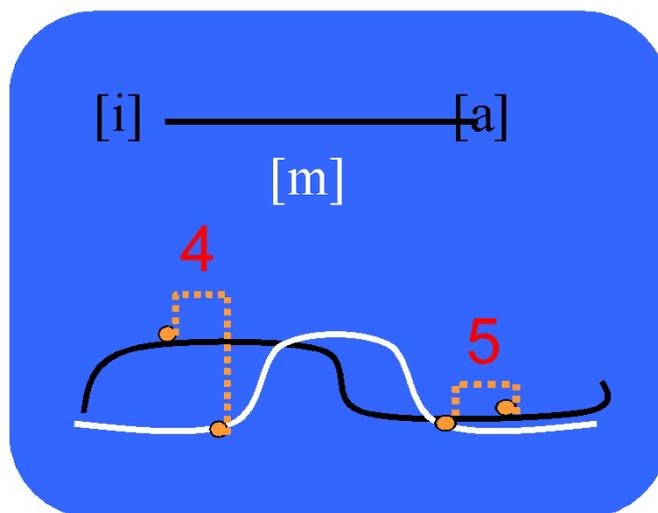


Fig 6. Schema delle misure effettuate per verificare il modello di coordinazione temporale C-V di Öhman (1967) applicato a “mim(m)a”. La linea nera si riferisce al movimento del dorso della lingua da [i] ad [a], mentre la linea bianca si riferisce al gesto dell’Apertura Labiale per la realizzazione della consonante [m]/[mm]. Vedi testo per la spiegazione.

4. RISULTATI

4.1 L’organizzazione temporale dei gesti fonetici

Per verificare se il fattore consonante singola/geminata fosse significativo per la coordinazione temporale tra consonate e vocale sono state eseguite una serie di analisi ANOVA a un fattore, separatamente per ciascun soggetto e per ciascuna velocità normale e massima.

Per quanto riguarda l’intervallo temporale tra il target articolatorio della prima vocale /a/ e il centro del ciclo consonantico /m(m)/ (vedi Fig. 5, n. 1) dovrebbe essere uguale per scempie e geminate, se per l’italiano fosse valido il modello di Öhman (1967). Infatti, in base ai risultati, non ci sono differenze significative, in entrambi i soggetti, sia a velocità normale, NE: [F(1,17)=1.992; p>0.05], NW: [F(1,17)=0.345; p>0.05]) che a velocità massima (NE: [F(1, 17)=0.003; p>0.05], NW: [F(1, 14)=0.205; p>0.05]).

D’altra parte, l’intervallo temporale tra il centro del ciclo consonantico /m(m)/ e il target articolatorio della seconda vocale /i/ (Fig. 5, n. 2) è maggiore per entrambi i soggetti nella produzione della geminata a velocità normale, sebbene una differenza significativa riguardi solo il soggetto NW ([F(1,17)=0.492; p>0.05], [F(1,17)=109.59; p<0.01]). A velocità massima l’intervallo è significativamente maggiore nelle geminate rispetto alla scempie solo per il soggetto NE ([F(1,17)=6.867; p<0.05]; NW: [F(1,14)=0.002; p>0.05]).

Relativamente all’intervallo temporale tra il target articolatorio per la prima vocale /i/ e il target della seconda vocale /a/ (vedi Fig.5, n. 3), che in base al modello di Öhman (1967), dovrebbe essere uguale per scempie e geminate, è invece significativamente maggiore per entrambi i soggetti nelle geminate prodotte a velocità normale (NE: [F(1,17)=4.435;

$p < 0.05$], NW: [F(1,17)=49.027; $p < 0.05$]) e per il soggetto NE anche a velocità massima ([F(1,17)=8.859; $p < 0.05$], NW: [F(1,14)=0.000; $p > 0.05$]) (v. fig. 7).

Sembrerebbe così che l'ipotesi di Smith, che prevede in base al modello di Öhman (1967) al variare della lunghezza fonologica l'invarianza degli intervalli temporali tra il punto centrale di massima costrizione consonantica e i target articolatori delle due vocali contestuali, non sia del tutto confermata. Questi risultati trovano spiegazione nel fatto che, a velocità di produzione normale, sia il gesto di chiusura come quello di apertura durano significativamente più a lungo nelle geminate rispetto alle scempie, (chiusura e apertura per NE: [F(1,18)=19.619; $p < 0.01$], [F(1,18)=32.794; $p < 0.01$]; chiusura e apertura per NW: [F(1,17)=12.405; $p < 0.01$], [F(1,17)=140.753; $p < 0.01$]). A velocità massima di produzione, solo il gesto di chiusura risulta significativamente maggiore nelle geminate rispetto alle scempie, e solo per il soggetto NE ([F(1,17)=34.297; $p < 0.01$]).

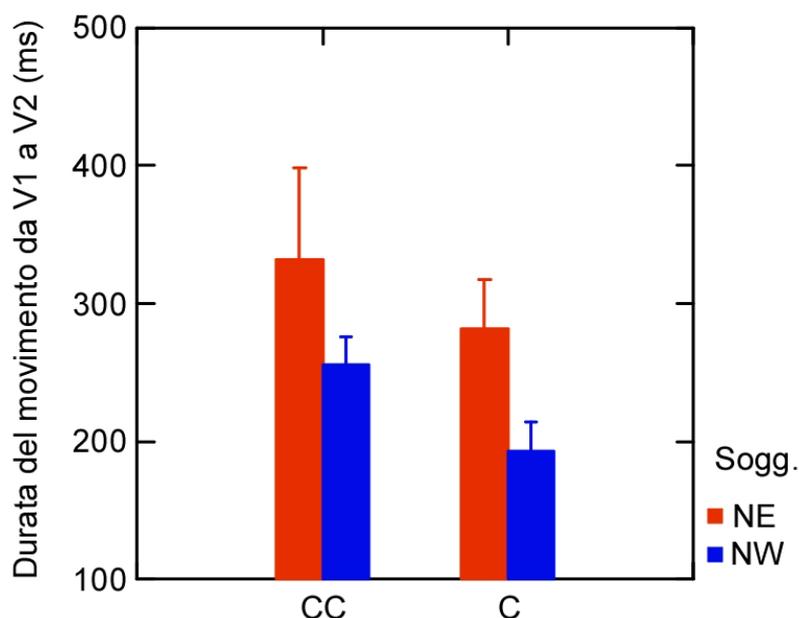


Fig. 7. Intervallo temporale dal target articolatorio della prima vocale /i/ a quello della seconda vocale /a/ per i due soggetti NE e NW e per la scempia (C) e geminata (CC)

La penultima misura, e cioè l'intervallo temporale tra l'istante relativo al target articolatorio del dorso della lingua per [i] e il momento iniziale della chiusura bilabiale per [m] e [mm] (v. Fig. 6, n. 4), di nuovo sembrerebbe soddisfare il modello di Öhman (1967), poiché effettivamente i risultati evidenziano che nelle geminate l'intervallo è significativamente più breve che nelle scempie per ambedue i soggetti a velocità normale (NE: [F(1,17)=4.450; $p < 0.05$], NW: [F(1, 17)= 15.994; $p < 0.05$], v. fig. 8), mentre non lo è a velocità massima (NE: [F(1, 17)=3.656; $p > 0.05$], NW: [F(1, 14)=0.058; $p > 0.05$]).

Invece l'intervallo speculare, che misura la durata tra la fine del gesto di chiusura bilabiale e l'istante relativo alla posizione del dorso della lingua per il target articolatorio [a] (vedi fig. 6, n. 5) non è significativamente diverso tra geminate e scempie, per nessun soggetto a nessuna velocità (normale e massima per NE: [F(1,17)=0.998; $p > 0.05$], [F(1,17)=3.791; $p > 0.05$]; per NW: [F(1,17)=1.151; $p > 0.05$], [F(1,14)=0.008; $p > 0.05$]).

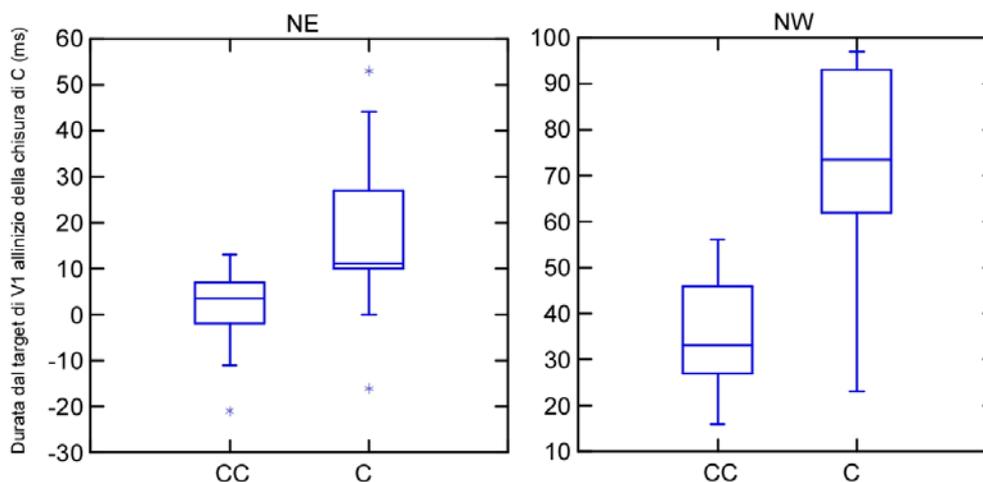


Fig.8. Intervallo temporale tra l'istante relativo al target articolatorio del dorso della lingua per [i] e il momento iniziale della chiusura bilabiale per [m] e [mm], per i due soggetti NE e NW per la scempia (C) e geminata (CC).

Riassumendo il senso di queste ultime due misure, si può dire che dall'analisi di queste due misure risulta che il gesto di chiusura per le geminate è anticipato rispetto a V1 (*ex hypothesis*), ma la fine del gesto di apertura non è ritardato rispetto a V2 (*contra hypothesis*).

Il quadro generale si qui delineato è confermato anche da altri risultati, ottenuti tramite analisi ispirate da un recente lavoro di Löfqvist (2006). Come nel caso di parlanti giapponesi, anche per i nostri soggetti l'inizio e la fine del movimento del dorso della lingua dalla prima alla seconda vocale occorrono rispettivamente prima e dopo il periodo di chiusura articolatoria della consonante, determinata acusticamente, e questo accade sia per le scempie che per le geminate. Solo per i nostri soggetti però, l'inizio della chiusura acustica occorre significativamente prima nelle geminate rispetto alle scempie (velocità normale e massima per NE: [F(1,17)=17.583; p<0.05], [F(1,17)=9.804; p<0.05] e per NW: [F(1,16)=11.069; p<0.05], [F(1,14)=0.605; p>0.05]). Il secondo intervallo speculare, quello che va dalla fine del periodo di chiusura consonantica, determinato acusticamente, e la fine del movimento della lingua per il target della seconda vocale, non rivela invece alcuna differenza.

Per concludere, la predizione di Smith (1995) di un'anticipazione del gesto consonantico nelle geminate rispetto alla prima vocale e di un suo ritardo rispetto alla seconda vocale è confermato solo nel primo caso.

4.2 L'ipotesi di Löfqvist (2006) del target gestuale "virtuale"

Ricordiamo brevemente che Löfqvist (2005) propose che i parlanti possono controllare la durata del periodo di chiusura degli articolatori, e quindi la distinzione acustica e percettiva legata alla differenza di durata tra consonate scempia e geminata, regolando la posizione spaziale del target consonantico.

Per testare l'ipotesi di Löfqvist (2005) abbiamo considerato la curva della velocità istantanea per l'apertura labiale in /mam(m)a/ e per il movimento verticale del dorso della

lingua per /l(l)a/ e /d(d)a/. Una serie di analisi ANOVA a un fattore sono state eseguite per verificare l'influenza della lunghezza consonantica su:

- il valore della curva di velocità al momento iniziale e al momento finale della chiusura consonantica, individuata dal punto di vista acustico rispettivamente come inizio e fine della fase di segnale a bassa energia per /m(m)/ /l(l)/ /d(d)/ (in pratica quando incomincia e termina il contatto tra l'articolatore attivo e quello passivo);
- la posizione temporale del picco di velocità rispetto a questi confini acustici.

I risultati, sia per i valore spaziali che temporali, sono riassunti in tabella 1 (in alto: velocità normale; in basso: velocità massima), dove S significa che la differenza tra scempie e geminate per quella misura è significativa ($p < 0.05$).

Per il gesto di apertura delle consonanti alveolari, la curva di velocità al rilascio della tenuta articolatoria ha un valore significativamente maggiore nelle geminate (v. fig. 9), e il suo picco di velocità è temporalmente più vicino all'istante di rilascio rispetto alle consonanti scempie (come accade per le bilabiali pronunciate da uno dei soggetti svedesi riportati in Löfqvist (2005)).

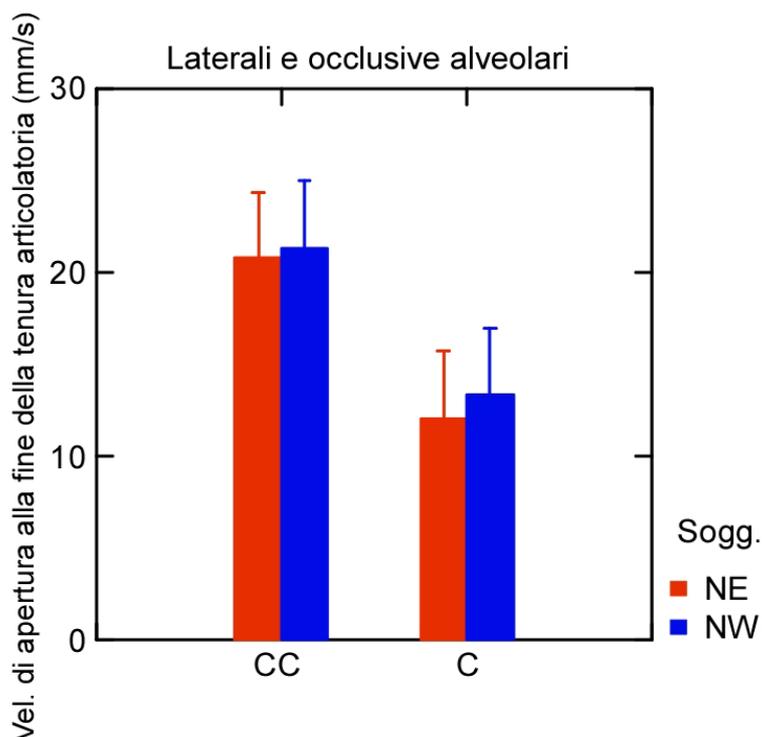


Fig. 9. Valore della velocità di apertura al momento finale della tenuta dell'occlusione alveolare, individuata dal punto di vista acustico, per i soggetti NE e NW, per la scempia (C) e geminata (CC).

A velocità massima, questa differenza tra geminate e scempie continua ad essere verificata solo per il soggetto NE, per tutti i fonemi considerati.

D'altra parte ci sono solo pochi risultati significativi per il gesto di chiusura a velocità normale e non riguardano mai lo stesso soggetto, mentre a velocità massima ci sono risultati significativo solo per il soggetto NE.

	Nas. Bilab.		Lat e Occl. Alv.	
	NE	NW	NE	NW
Chiusura	N : S	S : S*	N : N	N : N
Apertura	S : N	N : N	S : S	N : S

	Nas. Bilab.		Lat e Occl. Alv.	
	NE	NW	NE	NW
Chiusura	S : S	N : S	S : S	N : N
Apertura	S : S	S : N	S : S	N : N

Tab. 1. Risultati delle analisi ANOVA ad un fattore : velocità normale (in alto) e velocità massima (in basso), per entrambi i soggetti NE e NW. "S" significa che la differenza tra scempie e geminate è significativa ($p < 0.05$) per il valore di velocità all'inizio e alla fine del periodo di contatto articolatorio (a sinistra) e per la distanza temporale del suo picco massimo dai suddetti confini (a destra) nelle nasali bilabiali, e laterali e occlusive alveolari.

5. DISCUSSIONE E CONCLUSIONE

I risultati e le conclusioni di Smith (1995), che applica il modello di organizzazione temporale di Öhman (1967), non sono pienamente confermate dalle nostre analisi su due soggetti italiani che hanno pronunciato coppie di bisillabi differenziati per la lunghezza della consonante intervocalica. Il dato più forte a favore del modello di Öhman (1967) riguarda l'anticipazione dell'inizio del gesto consonantico per la geminata all'interno di V1. I dati in contrasto con tale modello però sono:

- il mancato ritardo della fine del gesto consonantico per la geminata all'interno di V2;
- l'assenza di invarianza per l'intervallo temporale tra i due target vocalici (V-to-V) . Viceversa esso varia significativamente al variare della lunghezza consonantica (*pro* Browman & Goldstein, *contra* Öhman);
- l'assenza di invarianza per gli intervalli tra il centro temporale dell'occlusione e il target articolatorio di V1 o V2. Essi non sono sempre costanti (come vorrebbe il modello di Öhman), soprattutto nel secondo caso.

Inoltre, a velocità massima, geminate e scempie mostrano poche differenze significative, e queste sono limitate a un singolo soggetto (NE).

Per quanto riguarda l'ipotesi di Löfqvist (2005) di un target articolatorio virtuale posizionato "più distante" per le geminate rispetto alle scempie, questa è parzialmente confermata. A favore ci sono i seguenti dati:

- a velocità normale, il gesto di apertura delle geminate alveolari ha una velocità significativamente maggiore, con il picco più vicino al rilascio.
- a velocità massima ciò è vero per ambedue i gesti di tutti i segmenti, ma solo per il soggetto NE.

A sfavore, la mancanza di differenze significative sistematiche per il gesto di chiusura, e a velocità massima, l'assenza di differenze anche per il gesto di apertura del soggetto NW.

6. BIBLIOGRAFIA

- Bertinetto, P.M. (1981), Strutture prosodiche dell'italiano, *Accademia della Crusca*, Firenze.
- Browman, C.P. & Goldstein, L.M. (1986), Towards an articulatory phonology, *Phonology Yearbook*, 3, 219-252.
- Esposito, A. & Di Benedetto M.G. (1999), Acoustical and perceptual study of gemination in Italian stops, *JASA*, 106, 4, 2051-2062.
- Gili Fivela, B., Zmarich, C. (2005), Italian Geminate under Speech Rate and Focalization Changes: Kinematics, Acoustic, and Perception Data, *InterSpeech 2005*, Lisbon, 2897-2900
- Gili Fivela, B., Zmarich, C., Perrier, P., Savariaux, C., Tisato, G. (2007), Acoustic and kinematic correlates of phonological length contrast in Italian consonants, in *Atti dell'International Conference of Phonetic Sciences (ICPhS'07)* ISBN:978-3-9811535-0-7, Saarbrücken, Germany, 6-10 agosto 2007, 469-472.
- Löfqvist, A. (2005), Lip kinematics in long and short stop and fricative consonants, *JASA*, 117, 2, 858-878.
- Löfqvist, A. (2006), Interarticulator programming: Effects of closure duration on lip and tongue coordination in Japanese, *JASA*, 120, 5, 2872-2883.
- Löfqvist, A. & Gracco, V. (1997), Lip and jaw kinematics in bilabial stop consonant production, *Journal of Speech, Language and Hearing Research*, 40, 877-893.
- Öhman, S.E.G. (1967), Numerical Model of Coarticulation, *JASA*, 41, 310-320.
- Smith, C.L. (1995), Prosodic patterns in the coordination of vowel and consonant gestures, *Papers in Laboratory Phonology IV, Phonology and phonetic evidence*, (B. Connell & A. Arvaniti, editors), CUP, 205-222.
- Tisato, G., Cosi, P., Somavilla, G., Zmarich C. (2007), Recent improvements in *InterFace* software: New tools for Articulatory Signal Processing, *Atti del III° Convegno Nazionale AISV* (V. Giordani, V. Bruseghini, P. Cosi, editors), Povo di Trento, 29-30 novembre-1 dicembre 2006, EDK Editore srl, Torriana (RN), 353-382.
- Zmarich, C., Gili Fivela, B. (2005), Consonanti scempie e geminate in italiano: studio cinematico e percettivo dell'articolazione bilabiale e labiodentale, in *Misura dei parametri, Atti del I° Convegno Nazionale dell'Associazione Italiana di Scienze della Voce (AISV)* (P. Cosi, editor), Padova, 2-4 dicembre, 2004, EDK Editore, Brescia, 429-448.
- Zmarich, C., Gili Fivela, B., Perrier, P., Savariaux, C., Tisato, G. (2007) Consonanti scempie e geminate in italiano: studio acustico e cinematico dell'articolazione linguale e bilabiale, in *Atti del Convegno Nazionale AISV, Scienze vocali e del linguaggio. Metodologie di valutazione e risorse linguistiche* (V. Giordani, V. Bruseghini, P. Cosi, editors), Povo di Trento, 29-30 novembre-1 dicembre 2006, EDK Editore srl, Torriana (RN), 151-163.