

FRANCESCA FONTANELLA, TOMMASO BALSEMIN, BARBARA GILI FIVELA,
PASCAL PERRIER, CHRISTOPHE SAVARIAUX, GRAZIANO TISATO,
CLAUDIO ZMARICH

L'organizzazione temporale delle geminate dell'italiano: uno studio di modellizzazione tramite la Fonologia Articolatoria

Modelling the temporal organization of geminates in Italian:
A study in Articulatory Phonology

Italian geminates are still a matter of debate about their phonological status and temporal organization with adjacent vowels. Although kinematic analyses can clarify what strategies are used to achieve length contrast, articulatory studies on Italian are scarce. The present work adopts the theoretical and methodological approach of *Articulatory Phonology* (Browman, Goldstein, 1989) to investigate the articulatory mechanisms involved in the gestural organization of Italian geminates, relying on the data of temporal intervals of consonantal versus vowel gestures, interpreted according to Gafos' (2002) formalization. The main objective is to test two alternative models of gestural synchronization, the "*Combined Vowel and Consonant*", proposed by *Articulatory Phonology*, and the "*Vowel-to-Vowel*", proposed by Öhman (1967). For this purpose, a "virtual geminate" was constructed by doubling the articulatory closure interval derived from the singleton and superimposing it on some contextual vowel *landmarks* (Gafos, 2002). The results of the different synchronizations of the virtual geminate according to the two models are then compared with those of the real geminate.

Keywords: bilabial geminates, Italian, Articulatory Phonology, synchronization VC(C)V.

1. Introduzione

Nel panorama internazionale, le geminate sono ancora oggi oggetto di discussione in ambito fonetico-fonologico e l'italiano è spesso citato come esempio nella letteratura internazionale per la pervasività della geminazione che lo caratterizza.

Una prima questione concerne lo statuto fonologico e vede contrapporsi due antitetiche ipotesi esplicative, una bisegmentale e eterosillabica (che Loporcaro, 1990, riconduce a Trubetzkoy, 1939) e una monosegmentale e tautosillabica (che Loporcaro, 1990, riconduce a Martinet, 1975). La prima proposta descrive le geminate come due segmenti identici posti l'uno in coda della sillaba precedente e l'altro in attacco della sillaba successiva. La seconda ipotesi vede le geminate come un unico segmento a livello fonologico, avente tratto [+lungo] e sillabificato interamente nell'*onset* della seconda sillaba.

Nel tentativo di supportare l'una o l'altra ipotesi, molti autori si sono serviti di "fatti" interni alla lingua, tra i quali l'osservazione delle regolarità distribuzionali e la derivazione dal latino. Tuttavia, i "fatti" di natura "esterna", relativi all'analisi acustica o cinematica, quantitativi e oggettivi, sono stati poco indagati, tanto più in italiano. In particolar modo, le analisi cinematiche potrebbero contribuire a spiegare quale delle quattro strategie per la realizzazione del contrasto di lunghezza descritte da Cho (2002) è quella realizzata preferenzialmente dal parlante italiano, e se ci sono varianti individuali o regionali. Collegata a ciò vi è poi la questione inerente al tipo di sincronizzazione con le vocali adiacenti.

L'analisi acustica ha aiutato a fare luce sui correlati della geminazione, il principale dei quali è la maggiore durata della fase di tenuta articolatoria. L'italiano non fa eccezione: diversi studi (Esposito, Di Benedetto, 1999; Mattei, Di Benedetto, 2001; Stevens, Hajek, 2004; Payne, 2005; Zmarich, Gili Fivela, 2005; Zmarich et al., 2006; Zmarich et al., 2007; Dipino, Celata, 2018; Celata, Meluzzi & Bertini, 2022) hanno mostrato che le geminate italiane sono lunghe circa il doppio della loro controparte scempia.

Invece, l'accorciamento della vocale precedente (V_1) il segmento raddoppiato, seppur noto (Esposito, Di Benedetto, 1999; Payne, 2005; Zmarich, Gili Fivela, 2005; Zmarich et al., 2006; Dipino, Celata, 2018; Mairano, De Iacovo, 2019; Di Benedetto, De Nardis, 2021; Celata, Meluzzi & Bertini, 2022), risulta più debole, in quanto influenzato da fattori quali l'accento (Bertinetto, 1981; Mairano, De Iacovo, 2019).

Circa l'analisi cinematica, lo studio di riferimento sul contrasto scempia-geminata è quello di Smith (1995), che si propone di verificare le ipotesi della Fonologia Articolatoria (Browman, Goldstein, 1989). Dato che il nostro lavoro utilizza lo stesso approccio teorico-metodologico, prima di descrivere i vari studi articolatori, si riporteranno i suoi principali assunti.

2. *Il contrasto scempia-geminata nella Fonologia Articolatoria*

2.1 Lo status delle geminate e la struttura della sillaba

La Fonologia Articolatoria (FA) è una teoria nata attorno agli anni '80, con i lavori pionieristici di E. Saltzman, P. Rubin, C. Browman, L. Goldstein (Browman, Goldstein, 1989).

Secondo FA, gli atomi di base della pianificazione fonologica e della produzione fonetica sono i gesti, azioni primitive degli articolatori presenti nel tratto vocale. Essi fungono sia da unità di contrasto che da unità d'azione (Browman, Goldstein, 1989), superando la tradizionale dicotomia tra fonologia e fonetica. Parte integrante di FA è il *Task Dynamic System* (TADA), esteso dal movimento degli arti, per il quale era stato originariamente concepito, a quello degli articolatori del parlato (Saltzman, 1986). TADA fornisce un meccanismo matematico per modellare il movimento dell'articolatore nello spazio e nel tempo, specificando come le coordinate spaziali del compito per realizzare un determinato bersaglio articolatorio (per es. la chiusura

labiale) vengano mappate sulle coordinate del tratto vocale e dell'articolatore, anche detto effettore (Poupplier, 2020). Così, ogni gesto è specificato da un aspetto spaziale intrinseco (i target delle variabili di tratto), ma anche da caratteristiche temporali intrinseche (es. *stiffness*, vedi oltre). Tali caratteristiche fanno sì che i gesti possano sovrapporsi (Fowler, 1980; Browman, Goldstein, 1989), ancorandosi in punti preferenziali della curva cinematica, denominati “punti di riferimento gestuale” (*landmark*), che sono: *onset* (o), l'inizio del gesto; *target* (t), l'istante in cui viene raggiunto il target articolatorio; *c-center* (cc), punto centrale tra raggiungimento del target e allontanamento dal target, ovvero il centro della fase di tenuta gestuale o “*plateau*”; *release* (r), momento in cui inizia l'allontanamento; *release offset* (roff), la fine del controllo attivo del gesto da parte dell'articolatore. Nel presente studio si farà riferimento al metodo di allineamento dei *landmark* secondo Gafos (2002).

Altro punto cardine di FA è la coordinazione inter-gestuale a livello della struttura sillabica, in quanto consonanti iniziali e finali si comportano diversamente. Quelle in attacco mostrano il cosiddetto effetto *c-center* (Browman, Goldstein, 1988): in inglese, ad esempio, l'aggiunta di ulteriori consonanti a inizio sillaba provoca un riaggiustamento della sincronizzazione dell'*onset*, tale che i gesti consonantici in attacco si coordinano con la vocale come fossero un blocco unitario. Lo stesso effetto non si manifesta nelle consonanti in coda: la prima consonante in coda ha una relazione così stabile con la vocale precedente, da non essere alterata dall'ipotetica aggiunta di ulteriori consonanti. La diversità di comportamento di *onset* e coda sillabici rappresenterebbe la manifestazione delle due fondamentali relazioni di fase (*phasing*) tra consonanti e vocali: “in-fase” e in “anti-fase” (Hall, 2010). La fisica dei sistemi biologici (Turvey, 1990) assume che il movimento di un gesto sia rappresentabile tramite il concetto di ciclo, che è assimilato a un angolo giro (360°). Nella relazione “in-fase” (o “fase zero”, con angolo di 0°) due gesti iniziano nello stesso istante. Nella sincronizzazione in “anti-fase” (cfr. *out-of-phase*), invece, il secondo gesto inizia quando il primo è a metà del tragitto, dando luogo a una temporizzazione spostata di 180° (come tamburellare le dita in alternanza). Circa la sincronizzazione consonante-vocale, si assume che la consonante in *onset* sia in una relazione di “fase” con la vocale, mentre la consonante in coda sia in una relazione di “anti-fase” con la vocale. Le consonanti, invece, affinché conservino la loro udibilità, devono coordinarsi tra loro in “anti-fase” (Browman, Goldstein, 1990).

Riassumendo, le relazioni di coordinazione sono tre: 1) C-V, relazione tra ciascun *onset* consonantico e la vocale successiva e che prevede che il *c-center* del gesto consonantico si sincronizzi con l'*onset* del gesto vocalico; 2) C-C, coordinazione tra le consonanti, sincronizzate in “anti-fase” tra loro (Browman, Goldstein, 2000)¹;

¹ Autori come Gafos, tuttavia, assumono che la coordinazione C-C sia linguo-specifica e non esista una relazione di default in tal senso (Gafos, 2002). In generale, le lingue presentano due modalità di transizione tra consonanti: le cosiddette “*close transition*” e “*open transition*”, le quali si rifletterebbero in una differente coordinazione. Nella “*close transition*” la costrizione articolatoria della seconda consonante si forma prima che la fase di tenuta per la prima consonante sia rilasciata (Catford, 1988; citato in Gafos, 2002). Nella “*open transition*”, invece, tra le consonanti facenti parte del nesso, emerge un vo-

3) V-C, relazione in anti-fase tra la vocale e la prima consonante post-vocalica occupante la coda sillabica. Le ipotetiche ulteriori consonanti poste in coda non presentano alcun tipo di coordinazione con la vocale, ma solo con la consonante che la precede (Browman, Goldstein, 1988).

Tuttavia, alcuni studi successivi (cfr. Marin, 2013, per il rumeno; Pastätter, Pouplier, 2015, per il polacco) evidenziano come la sincronizzazione C-V non è spiegabile in modo uniforme mediante la teoria di organizzazione della sillaba. Infatti, la coordinazione del *c-center* con la vocale dipende dalla resistenza alla coarticolazione, che varia in base allo specifico segmento consonantico presente in attacco. Tale peculiarità non è specificata da Browman, Goldstein (2000).

Fatte tali premesse, due gesti che a parità di ampiezza presentano una diversa durata, possono essere regolati da meccanismi che agiscono sulle caratteristiche intragestuali (es. *stiffness*, o costante di rigidità) o sulle relazioni intergestuali di *phasing*.

2.2 Indagini cinematiche relative al contrasto scempia geminata

Lo studio di Smith (1995) è considerato il capostipite delle successive indagini sulle strategie cinematiche implicate nella geminazione italiana. La ricercatrice confronta la geminazione dell'italiano con quella del giapponese, analizzando parole con struttura mVC(C)V (con V = [a, i] e C = [p, t, m, n]), anche se poi l'analisi si limita alle parole con struttura miC(C)a. I risultati mostrano che l'italiano si conforma al modello "*Vowel-to-Vowel timing*" (*V-to-V*) di Öhman (1967; citato in Smith, 1995), mentre il giapponese obbedisce al "*combined vowel and consonant timing*" (*V-to-C*) di Browman, Goldstein (1990). Il modello *V-to-V*, che Smith (1995) associa a lingue *stress-timed* (come l'inglese) e *syllable-timed* (come l'italiano), assume che l'articolazione delle sillabe si basi sulla produzione delle vocali interpretate come successione di cicli articolatori. Le consonanti, invece, sono articolate separatamente come modulazione del ciclo articolatorio delle vocali. Dunque, le consonanti sono irrilevanti per la sequenziazione vocalica che, quindi, non dovrebbe essere influenzata né dalla durata consonantica né da loro altre caratteristiche (Smith, 1995). Contrariamente, nel modello *V-to-C*, che Smith attribuisce alle lingue *mora-timed* (come il giapponese), vocali e consonanti sono reciprocamente coordinate. Pertanto, anche le consonanti possono influenzare la temporizzazione vocalica, cosicché una maggiore durata della consonante mediana corrisponderebbe a un allungamento da vocale a vocale (Smith, 1995). Nella geminazione italiana il ciclo vocalico sarebbe mantenuto costante grazie a un'anticipazione del gesto di chiusura consonantica a livello di V₁ e un ritardo nel gesto di rilascio della stessa consonante a livello di V₂, dunque grazie a una maggiore sovrapposizione. Inoltre, Smith (1995) evidenzia come le geminate siano caratterizzate da una maggiore durata del gesto di chiusura e della fase di tenuta articolatoria: la prima dovuta a una minore rigidità (*stiffness*); la seconda conseguente a un periodo di attivazione più lungo. Infine,

coide di transizione, il quale permette un periodo di transizione tra i due gesti utile affinché entrambi siano udibili da parte dell'ascoltatore.

Smith (1995) rileva un solo picco di ampiezza nelle geminate, corrispondente al target, come nelle scempie, sottolineando con questo la problematicità di un modello bisegmentale/bigestuale.

Successivamente, Zmarich, Gili Fivela (2005) hanno registrato parole e pseudo-parole con consonanti alternativamente scempie e geminate ([f, v, p, b, m]) e struttura CaCCa, e stimoli contenenti nessi consonantici (/mf/, /mb/ e /mp/), prodotti da due parlanti originari dell'Italia settentrionale; i risultati cinematici, per i target prodotti a velocità normale e focalizzazione ampia, mostrano differenze significative tra scempie e geminate: le geminate somigliano ai nessi consonantici, soprattutto a livello del gesto di apertura del labbro inferiore, il quale presenta durata e ampiezza maggiori, una durata maggiore per il raggiungimento del picco di velocità (*time-to-peak*) e minore rigidità (*stiffness*). Tale similarità ha condotto Zmarich, Gili Fivela (2005) a ipotizzare l'eterosillabicità delle geminate. La differenza maggiore riscontrata tra scempie e geminate, soprattutto a livello del gesto di apertura labiale verso la seconda vocale, ha confutato l'ipotesi di Smith (1995), secondo la quale l'occlusione labiale è rappresentata da un gesto in cui le fasi di chiusura e apertura sono simili. Gli studiosi, inoltre, confermano parzialmente l'appartenenza dell'italiano al modello *V-to-V* di Öhman (1967), suggerendo invece come la sincronizzazione temporale sia soggetta a variazione interindividuale.

Lo stesso anno, Gili Fivela, Zmarich (2005) hanno analizzato gli stimoli di Zmarich, Gili Fivela (2005) prodotti a velocità d'eloquio rapida e nelle condizioni prosodiche "non in focus" e in "focus contrastivo". I dati con focalizzazione contrastiva confermano quanto evidenziato da Zmarich, Gili Fivela (2005): le geminate assomigliano ai nessi consonantici eterosillabici, soprattutto circa le caratteristiche del gesto di apertura del labbro inferiore. Con l'aumento della velocità d'eloquio, le differenze cinematiche, a differenza di quelle acustiche, vengono preservate e ciò conduce a ipotizzare che i correlati cinematici della geminazione siano più stabili.

Zmarich et al. (2006) hanno studiato, mediante l'*Electromagnetic Midsagittal Articulometer* 2D (EMA; Hoole, Nguyen, 1999), parole e pseudo-parole di struttura 'CVC(C)V contrastanti per lunghezza (con V=[a] e C=occlusive sonore dentali, velari e laterali alveolari), e parole e non parole contenenti gruppi consonantici (del tipo /ld/ o /dl/), prodotti da due locutrici di italiano settentrionale. Le analisi, parimenti a quanto rilevato in Zmarich, Gili Fivela (2005), mostrano che le geminate sono modellate da due gesti, chiusura e apertura, e che somigliano maggiormente ai gruppi consonantici eterosillabici in termini di maggiore durata, maggiore ampiezza, minore rigidità (*stiffness*), maggiore tempo di raggiungimento del picco di velocità (*time-to-peak*), cioè minore accelerazione, maggiore velocità nel gesto di chiusura e minore velocità del gesto di apertura. La somiglianza è maggiore a livello del gesto di apertura rispetto a quello di chiusura, evidenziando la rilevanza di tale gesto nel contrasto di lunghezza consonantica.

Un'indagine elettropalatografica (EPG) di Payne (2006) su coppie minime scempia-geminata con struttura "aC(C)a" prodotte da una locutrice romana, ha evidenziato una conformazione più piatta della superficie della lingua nelle

geminate e più concava nelle scempie. Tali evidenze supportano le ipotesi di Payne (2005), secondo cui il contrasto di lunghezza in italiano non è dettato da una mera differenza di durata temporale della fase di occlusione, ma anche da una differenza in termini spaziali.

Zmarich et al. (2007) e Gili Fivela et al. (2007) hanno analizzato gli stimoli di Zmarich, Gili Fivela (2005) per la consonante /m/ in contesto /i-/a/. Alla velocità di eloquio normale, il ciclo vocalico varia significativamente in base alla lunghezza consonantica, diversamente dal modello *V-to-V* e similmente a quello *V-to-C*, pur mantenendo l'anticipazione del gesto consonantico, coerente col modello *V-to-V*, trovata anche da Smith (1992). Gli studiosi concludono che le geminate italiane sembrano conformarsi a un "modello ibrido", composto nella prima parte dal *V-to-V* e nella seconda parte dal *V-to-C*.

Uno studio con EMA 2D di Zmarich et al. (2011) sulle strategie articolatorie per la realizzazione del contrasto di lunghezza consonantica nella coppia minima "*mima/mimma*", confuta l'ipotesi del modello *V-to-V* sia alla velocità di elocuzione normale che a quella rapida: il ciclo vocalico varia significativamente tra scempie e geminate. Tuttavia, la predizione di Smith (1995) circa l'anticipazione del gesto consonantico della geminata nella vocale precedente riceve una nuova conferma. Alla velocità d'eloquio rapida le differenze tra scempie e geminate si riducono.

Indagini relative alle geminate sono state svolte anche utilizzando altri metodi e strumenti di misurazione. Hagedorn, Proctor & Goldstein (2011), mediante rtMRI, hanno indagato coppie minime scempia-geminata, costituite da occlusive, affricate e sonoranti, prodotte da un locutore di Roma. I risultati mostrano una differenza significativa nel picco di velocità massima (*peak intensity*) nelle geminate occlusive bilabiali, coronali laterali-nasali-occlusive; nelle occlusive coronali, tuttavia, la differenza non è significativa. Circa la cinematica gestuale, l'intera durata della consonante (dall'inizio al rilascio) è maggiore nelle geminate.

Dipino, Celata (2018) hanno effettuato un'ispezione dei profili sagittali della lingua per le occlusive alveolari, registrando cinque locutori della varietà di toscano tramite UTI (cfr. Stone, 2005). Coerentemente con Zmarich et al. (2007), le geminate sono prodotte con un gesto linguale più ampio.

Celata, Meluzzi & Bertini (2022) hanno confrontato per mezzo dell'UTI i gruppi consonantici tautosillabici ed eterosillabici italiani per comprendere se le caratteristiche temporali e cinematiche proprie del contrasto scempia-geminata valevano anche per il contrasto tra due tipi di gruppi consonantici, partendo dal presupposto che in entrambi i casi ci fosse un'opposizione tra sillaba aperta e chiusa. I due effetti cinematici della geminazione, ovvero l'aumento della velocità di chiusura e l'aumento dell'anticipazione temporale dell'inizio del gesto di chiusura nella vocale precedente, sono stati confermati in tutti i tipi di consonanti e in tutti i contesti all'interno del contrasto scempia-geminata, ma non in quello inerente alle due tipologie di gruppi consonantici. Anche il correlato acustico corrispondente all'accorciamento della vocale precedente la consonante in coda sillabica si è mostrato meno pregnante nel contrasto tra le due tipologie di gruppi consonantici.

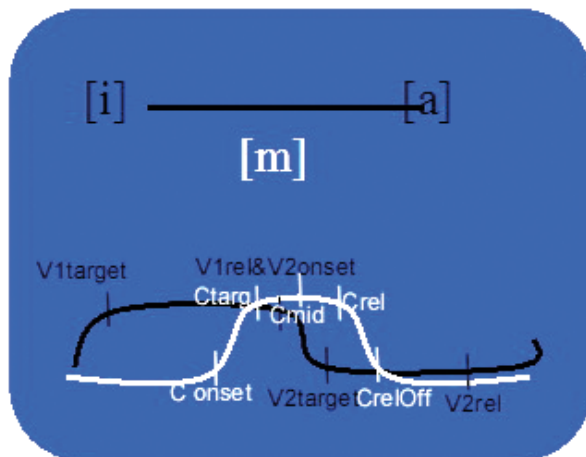
Tali risultati hanno evidenziato come le caratteristiche cinematiche e temporali che contraddistinguono il contrasto scempia-geminata sono proprie del contrasto in sé e non della specifica struttura sillabica.

Infine, Burroni et al. (2024) hanno indagato le proprietà cinematiche delle geminate italiane mediante EMA (3D, Carstens AG501), utilizzando pseudo-parole bisillabiche “VC(:)V” (con C= [p], [m] e V=[i], [a]) a 5 diverse velocità di elocuzione. Lo studio ha confermato la maggiore durata nella fase di chiusura per le geminate. Circa le proprietà cinematiche, le geminate sono prodotte, a prescindere dalla velocità di elocuzione, con movimenti articolatori più ampi, velocità massima leggermente minore e target più estremi (come in Löfqvist, 2005). Burroni et al. sostengono che le geminate non siano meramente delle scempie con intervalli di attivazione più lunghi (cfr. Gafos e Goldstein, 2012), bensì una specifica categoria gestuale. Una prova di ciò deriva anche dalla differente temporizzazione delle geminate con le vocali adiacenti: viene confermata l’anticipazione del gesto di chiusura della consonante in caso di geminazione rispetto a V_1 (come in Smith, 1995, e come in tutti gli autori fin qui considerati); tuttavia, l’intervallo vocalico V_1 - V_2 , differentemente dal modello *V-to-V*, risulta minore in caso di interposizione di una geminata. Inoltre, gli studiosi mostrano come anche il gesto per V_2 inizia prima se è presente una consonante raddoppiata, differentemente da quanto emerso in Smith (1995). Tali contraddizioni, rispetto a studi citati in precedenza, possono derivare da differenti scelte nel posizionamento dei vari *landmark*.

3. Obiettivi e ipotesi

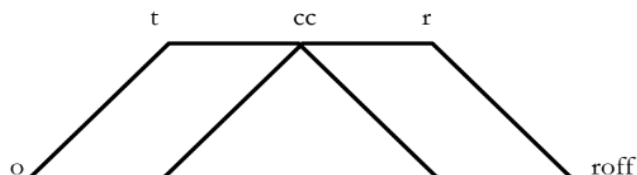
Il presente studio si pone diversi obiettivi inerenti alla sincronizzazione gestuale delle consonanti scempie (VCV) e geminate (VCCV), utilizzando i dati raccolti e parzialmente elaborati da Zmarich et al. (2011). I dati ivi trattati sono esclusivamente di tipo temporale: altre variabili cinematiche (ampiezza, *stiffness*, *time-to-peak* etc.) non sono state indagate. I dati analizzati riguardano la velocità d’eloquio “normale” (N, *normal*). Prima di ragionare sugli intervalli, è stata modificata la Fig.1 che compariva in Zmarich et al. (2011), adottando la terminologia dei *landmark* di Gafos (2002).

Figura 1 - Landmark della traiettoria cinematica di “mim(m)a” secondo la formalizzazione di Gafos (2002) (adattata da Zmarich et al., 2011). Traiettoria linguale (nella dimensione verticale) per la prima e seconda vocale (nero); apertura labiale (LA) per la consonante (bianco). Si noti che “Cmid”=“Ccent”



Si noti come la traiettoria del dorso linguale per le vocali (in nero, Fig. 1) non possieda tutti i *landmark* della formalizzazione di Gafos (2002, Fig. 2). Inoltre, si osservi come il rilascio della prima vocale (*V1rel*) coincide con l'inizio della seconda vocale (*V2ons*).

Figura 2 - Sovrapposizione di due curve gestuali in “close transition” come da formalizzazione di Gafos (2002)



Il primo obiettivo del lavoro è quello di testare due modelli alternativi di sincronizzazione gestuale, il “combined vowel and consonant timing” (*V-to-C*, Browman, Goldstein, 1990) e il “Vowel-to-Vowel timing” (*V-to-V*, Öhman, 1967).

Di seguito riportiamo gli intervalli ritenuti cruciali per il confronto, nonché le nostre predizioni a riguardo, aiutandoci con le Figg. 3 e 4:

- *V1targ_Cons* (indicato con 1 nelle Figg. 3 e 4) di VCCV è diverso ($\neq$) dallo stesso intervallo in VCV per entrambi i modelli. Per il *V-to-C* tale intervallo dovrebbe variare in caso di geminazione, in quanto l'aggiunta di ulteriori consonanti modifica la sincronizzazione vocalica. Infatti, Browman, Goldstein (1990) prevedono che in VCV la consonante sia sincronizzata con la seconda vocale, mentre in VC1C2V, C1 è sincronizzata con la prima vocale. Invece,

per il modello *V-to-V*, l'intervallo *V1targ_Cons* differisce tra VCV e VCCV, in quanto mantenendo il ciclo vocalico inalterato, (cioè mantenendo invariato il valore temporale corrispondente a *V1onset*, si assiste ad un'anticipazione del gesto consonantico per consentire la sua espansione, cioè *Ctarget* si avvicina a *V1*).

- *V1targ_Ccent* (indicato con 2 nelle Figg. 3 e 4) di VCCV è diverso ($\neq$) dallo stesso intervallo in VCV secondo *V-to-C*, in quanto risente della durata consonantica e della sincronizzazione in “anti-fase” della prima parte della consonante. Nella scempia, invece, la prima vocale non ha alcuna relazione di fase con la vocale precedente. Nel modello *V-to-V* questo intervallo si mantiene uguale ($=$).
- *Ccent_V2rel* (indicato col 3 nelle Figg. 3 e 4) di VCCV è simile ($=$) allo stesso intervallo in VCV secondo *V-to-C*, in quanto lo spostamento verso destra di *Creloff*, secondo Browman, Goldstein (1990) è compensato parzialmente dallo spostamento di *V2* sempre nella stessa direzione. Per motivi diversi questo intervallo resta invariato ($=$) anche nel modello *V-to-V*.

Figura 3 - Schema degli intervalli temporali indagati per verificare il modello di coordinazione *V-to-V* di Öhman (1967) applicato a “mim(m)a”. La curva nera rappresenta il movimento del dorso della lingua da [i] ad [a]; la linea bianca simula il gesto dell'Apertura Labiale per la realizzazione della consonante [m]/[mm]. Vedi testo per la spiegazione sopra=singola; sotto=geminata

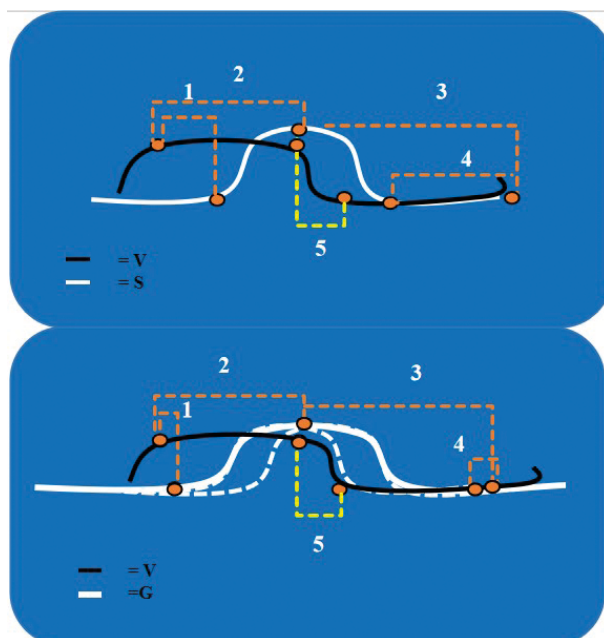
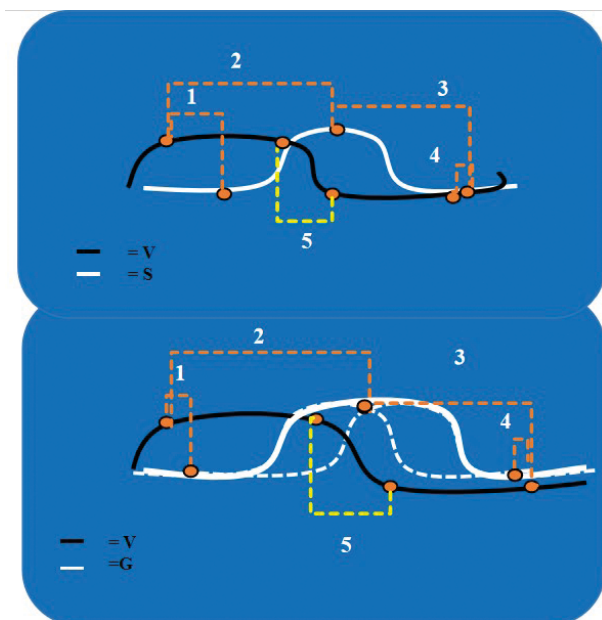


Figura 4 - Schema degli intervalli temporali indagati per verificare il modello di coordinazione V-to-C di Browman e Goldstein (1990) applicato a "mim(m)a". La curva nera rappresenta il movimento del dorso della lingua da [i] ad [a]; la linea bianca simula il gesto dell'Apertura Labiale per la realizzazione della consonante [m]/[mm]. Vedi testo per la spiegazione sopra=singola; sotto=geminata



- $Crel_{off_V2rel}$ di VCCV (indicato col 4 nelle Figg. 3 e 4) è simile (=) allo stesso intervallo in VCV secondo Browman, Goldstein (1990), in quanto V2 subirà un ritardo gestuale causato dalla maggiore lunghezza consonantica, compensata dallo spostamento a destra di V2. Invece, nel caso del modello di Öhman (1967), il suddetto intervallo sarà diverso ($\neq$) dallo stesso intervallo in VCV. Ciò si spiega col fatto che lo spostamento a destra di $Crel$ non influisce su $V2rel$, che rimane fisso.
- $V1rel_V2targ$ (indicato col 5 nelle Figg. 3 e 4) non era stato calcolato da Zmarich et al. (2011). L'intervallo rappresenta la transizione dalla fine di V1 all'inizio di V2. Basandosi su studi precedenti di Zmarich et al. (2005, 2007, 2011), si ipotizza che $V1rel_V2targ$ in VCCV sia ($\neq$) dallo stesso intervallo in VCV secondo V-to-C, ovvero maggiore nel primo caso, in quanto risente della lunghezza consonantica; contrariamente, secondo il modello di Öhman (1967) tale intervallo, relativo al ciclo vocalico, rimarrà invariato tra scempia e geminata.

Un ulteriore obiettivo riguarda lo studio dei singoli intervalli consonantici ricavabili dai 4 *landmark* del ciclo consonantico proposti da Gafos (2002, Fig. 2), ossia $Cons_Ctarg$, $Ctarg_Crel$ e $Crel_Crel_{off}$ (Fig. 1). Tale analisi non è funzionale alla verifica di alcun modello, ma è finalizzata a comprendere se il contrasto di lunghezza in italiano sia motivato dal prolungamento di un unico intervallo temporale, per

esempio la fase di tenuta dell'occlusione, che è il correlato acustico pregnante per la geminazione. Si potrebbe ipotizzare che il prolungamento dell'intervallo di tenuta sia ricavato tramite un'anticipazione del gesto consonantico nella geminazione, nonché un ritardo del suo rilascio, i cui effetti acustici dovrebbero consistere in una riduzione della durata delle due vocali contestuali come predetto dal modello di Öhman (1967), richiamato da Smith (1995) e da Zmarich et al. (2007).

Un'ultima modalità di verifica e confronto dei modelli cinematici V-to-V e V-to-C è consistita nella creazione di una "geminata virtuale" (GV, si veda § 4.3).

4. Metodi

4.1 Corpus e partecipanti

Il corpus sperimentale è stato mutuato da Zmarich et al. (2005) ed è formato dalla coppia minima scempia-geminata "*mima/mimma*". Le parole target erano inserite entro la frase cornice "*richiama mim(m)a malamente*". Le frasi erano state randomizzate e presentate a gruppi di tre su pannelli posti a circa due metri di distanza dai soggetti. Erano state previste dieci ripetizioni per frase a velocità normale e altre dieci a velocità sostenuta, sebbene nel presente studio vengano riportati solo i risultati inerenti alla velocità normale.

I quattro soggetti sono gli stessi dell'esperimento descritto da Zmarich et al. (2005), Gili Fivela et al. (2007) e Zmarich et al. (2011). La locutrice (AG) appartiene a una varietà italiana del nord-est; due soggetti, una femmina (BG) e un maschio (MP), parlano un italiano regionale del nord-ovest, mentre (FC) proviene da una varietà del centro Italia. Tutti i locutori si erano dichiarati esenti da qualunque disturbo di linguaggio e/o voce, presente o pregresso.

4.2 Acquisizione dei materiali e misure

Il segnale audio è stato registrato con due dispositivi distinti, uno integrato all'hardware di EMA (vedi oltre) e l'altro esterno. Il segnale analizzato è quello acquisito col dispositivo esterno, di maggior qualità (44 kHz e 16 bit), costituito da un DAT collegato a un microfono professionale. È stato utilizzato il software PRAAT (<https://www.fon.hum.uva.nl/praat/>) per l'analisi del segnale acustico e per la segmentazione manuale delle parole entro i loro confini. Nello specifico, la sequenza VC(C)V è stata segmentata e misurata mediante lo studio del sonogramma a banda larga. Nel caso delle geminate non si sono posti confini acustici interni alla sequenza consonantica.

Per la raccolta dei dati cinematici, è stato utilizzato l'*Electro Magnetic Articulograph* o EMA AG200 (© Carstens), nella versione 2D all'epoca in uso presso il laboratorio ICP-CNRF (ora GIPSA lab) a Grenoble, in Francia. I sensori erano stati posizionati, tramite l'uso di colla tissutale, nelle seguenti posizioni (con riferimento al piano medio-sagittale della testa): due di essi, utili per compensare i movimenti della testa, erano stati collocati sul ponte nasale (lo spazio tra le

sopraciglia) e sul prolabio (solco sotto-nasale); due erano stati posti al centro del labbro inferiore e superiore; quattro erano stati posizionati sulla superficie linguale a distanza l'uno dall'altro di circa un centimetro, partendo da un centimetro subito dopo l'apice linguale e proseguendo verso la radice.

Per le consonanti bilabiali si sono utilizzate le misurazioni dei sensori incollati al centro delle labbra, mediante i quali è stata rilevata l'apertura labiale (*Lip Aperture*), ovvero la distanza verticale. Per ottenere dati sulla cinematica delle vocali, si è fatto ricorso al sensore posto sul dorso della lingua, posizionato a circa 3.5 cm dall'apice della lingua.

Le misurazioni cinematiche sono state eseguite in una fase successiva alla segmentazione e all'etichettatura del segnale acustico e sono state allineate a queste ultime. Per l'elaborazione dei dati cinematici si è impiegato il software *Interface* (cfr. Tisato, Così, Somnavilla, Zmarich, 2007). Similmente a Smith (1995), per ottenere l'inizio e la fine di ogni gesto si sono utilizzati i punti in cui la velocità dell'articolatore in esame rispettivamente raggiungeva o diminuiva del 10 % rispetto al suo picco massimo.

Alle misure acustiche e cinematiche, Zmarich et al. (2011) avevano affiancato anche un test di percezione, volto a verificare se i locutori dello studio avessero prodotto adeguatamente le consonanti geminate. Infatti, in alcune varietà italiane, nella fattispecie settentrionali, il contrasto scempia-geminata è segnalato con una durata di occlusione minore (degeminazione settentrionale) e percettivamente può risultare indebolito o assente. Con il software *Perceval* (<https://blog.bitergia.com/2018/05/01/perceval-software-project-data-at-your-will/>) cinque ascoltatori del nord Italia e cinque del sud Italia avevano ascoltato le tracce audio contenenti le frasi cornice e avevano giudicato se le parole target al loro interno fossero geminate o meno. Le produzioni qui analizzate sono quelle risultate statisticamente prototipiche di ciascuna delle due categorie di lunghezza, e tra loro contrastanti (per i dettagli della verifica percettiva, si rimanda agli articoli citati).

La procedura sperimentale effettuata ex novo, univoca per tutti gli obiettivi sopra citati, è consistita nel calcolo di intervalli temporali ritenuti utili alle verifiche prefissate. Gli intervalli temporali sono stati ricavati sottraendo o sommando i vari *landmark* gestuali già calcolati da Zmarich et al. (2011), e altri ricavati da *landmark* "virtuali", espressamente ricavati per questo lavoro. Gli intervalli ritenuti utili alle verifiche inerenti al confronto tra i modelli *V-to-V* e *V-to-C* si sono rivelati essere: *V1targ_Cons*, *V1targ_Ccent*, *Ccent_V2rel*, *Crelloff_V2rel*, *V1rel_V2targ*. Gli intervalli utilizzati per testare il loro ruolo nel contrasto scempia-geminata sono stati: *Cons_Ctarg*, *Ctarg_Crel* e *Crel_Crelloff*. Questi intervalli sono poi stati organizzati in un database *Excel*, utili a effettuare le successive analisi statistiche.

4.3 Costruzione della "geminata virtuale"

Nella parte che segue descriveremo le modalità con cui abbiamo costruito la cosiddetta "geminata virtuale" (GV). In primis, si è raddoppiata la curva della scempia partendo dai dati reali ottenuti da Zmarich et al. (2011), sovrapponendo

successivamente le due curve ottenute in modalità *close transition*. Tale scelta si attiene a Gafos (2002), il quale critica l'idea della FA inerente a rapporti di fase consonantica di tipo universale, assumendo che questi siano linguo-specifici. Nel caso dell'italiano non vi sono prove a favore di una transizione aperta in caso di geminazione (consideriamo qui come occasionale la presenza di due *bursts*, indici acustici del rilascio di un'occlusione, riscontrati da Di Benedetto et al., 2021 nelle loro registrazioni). Per ottenere la sovrapposizione consonantica in "*close transition*", il *C1rel* è stato allineato con il *C2targ*, creando il *CcenterGV* della geminata virtuale (*cc* nella Fig. 2).

Dapprima si sono ricavati i vari *landmark* di GV: l'*onset* e il *target* non sono stati ricavati ex novo, in quanto corrispondenti a quelli della consonante scempia "reale"; anche per il *CcenterGV* non è stata necessaria alcuna derivazione, in quanto corrispondeva al *Crel* della scempia; il *CrelGV* in *close transition* è stato computato partendo dal *Crel* della scempia e aggiungendovi l'intervallo *Ctarg_Crel* della stessa; per ottenere il *CreloffGV* in *close transition*; infine, l'intervallo *Crel_Creloff* della scempia è stato aggiunto al *CrelGV*.

Successivamente, l'organizzazione temporale della GV è stata costruita in modo tale da simulare l'organizzazione temporale di ciascuno dei due modelli di sincronizzazione, *V-to-V* e *V-to-C*, al fine di confrontare statisticamente i dati simulati di alcuni intervalli critici con quelli ottenuti dalle geminate effettivamente prodotte (cfr. Zmarich et al., 2011), e verificare così l'adeguatezza dei modelli. A tale scopo si sono allineati i *landmark* della GV ottenuta a quelli delle vocali contestuali alla consonante scempia.

A questo scopo abbiamo utilizzato il modello *V-to-V* in due modi. Nel primo modo, lavorando sulle matrici di *Excel*, l'espansione del ciclo consonantico della geminata rispetto alla scempia è stata simulata facendo coincidere il *CcentGV* al *Ccent* della scempia, ed espandendo i *landmark* di GV a sinistra e a destra di un intervallo temporale corrispondente a metà della durata di occlusione articolatoria *Ctarget_Crel* della scempia. In termini pratici questo significa che abbiamo sottratto il valore temporale di tale intervallo ai *landmark* consonantici a sinistra del *CcentGV* (anticipando così la loro occorrenza temporale rispetto ai corrispondenti *landmark* della scempia), ottenendo *ConsGV* e *CtargGV*. Parallelamente, abbiamo aggiunto il valore temporale di tale intervallo a tutti i *landmark* a destra del *CcentGV* (posticipando temporalmente la loro occorrenza rispetto ai corrispondenti *landmark* della scempia), ottenendo *CrelGV* e *CreloffGV*. Particolare importante, abbiamo lasciato l'organizzazione temporale del ciclo vocalico inalterata, e non abbiamo fatto riferimento ad alcun *landmark* della vocale per simulare l'organizzazione temporale di GV.

Con la seconda modalità di simulazione di GV in base al modello *V-to-V*, assumiamo che il *Ccent* sia ancorato al *V2onset*, adottando la proposta di sincronizzazione che Browman, Goldstein (1988, 2000) avanzano per la sincronizzazione della consonante iniziale nella sillaba CV. Pur riconoscendo che si tratta di una forzatura, può essere però un'ipotesi di lavoro interessante nel senso

che questo tipo di sincronizzazione “cala nella realtà” la proposta fonologica di una consonante lunga sillabificata in *onset* (V.CCV), avanzata per primo da Martinet (1975). Pertanto, si è fatto corrispondere il *CcenterGV* (coincidente col *Crel* della scempia) al *V2ons* del ciclo vocalico VCV; successivamente è stata calcolata la differenza tra il *Ccenter* originario e il nuovo *CcenterGV* (ovvero, *V2ons*), al fine di sottrarre questo intervallo a ogni *landmark* della scempia e ottenere i nuovi punti della GV. In questo modo, togliendo la differenza (*Ccenter-CcenterGV*) si sono ottenuti il *CtargGV*, *CrelGV*, *CreloffGV*. Poiché il *Ccent* della scempia occorreva diversi ms dopo il *V2onset*, dove è stato ancorato il *Ccent* della GV, l'effetto prodotto è consistito in uno spostamento globale a sinistra dell'intera configurazione della GV.

Di seguito si è testato il modello *V-to-C*, tenendo conto dei vari rapporti di fase, e partendo dal presupposto che, secondo quest'ultimo, la quantità consonantica modifica il ciclo vocalico (allungandolo). Dunque, sono state simulate le diverse sincronizzazioni tra consonante e vocali contestuali: il *CtargGV* è stato allineato al *VIrel* della scempia, riproducendo così la temporizzazione in “anti-fase” tra la consonante in coda sillabica e la vocale precedente; successivamente, è stata calcolata la differenza tra il *Ctarget* della scempia e il *CtargGV*; tale differenza, poi, è stata sottratta ai *landmark* della scempia per ottenere gli ancoraggi della GV (*ConsGV*, *CcentGV*, *CrelGV* e *CreloffGV*); successivamente, si è fatto coincidere il *V2onsetGV* con il *CcenterGV*, simulando la sincronizzazione “in-fase” tra la consonante in *onset* e la vocale successiva, nel senso che *CcenterGV* impone che *V2onsetGV* sincronizzi con esso; di seguito si è calcolata la differenza tra il *V2onsGV* e il *V2onset* della scempia, in modo che anche i *landmark* successivi della seconda vocale fossero spostati a destra, risentendo della lunghezza consonantica; si sono quindi calcolati i *landmark* *V2targGV* e *V2relGV*, sottraendo la differenza di cui sopra ai *landmark* *V2targ* e *V2rel* della scempia.

Con un'ultima simulazione abbiamo tentato di riprodurre l'ipotesi avanzata da Zmarich et al. (2007) e Gili Fivela et al. (2007), secondo cui le geminate italiane si conformano a un “modello ibrido”, che possa spiegare sia la riduzione della durata temporale che sia ha in V_1 di V_1CCV_2 , sia il sostanziale mantenimento in V_2 . Per questo motivo, la serie di sincronizzazioni è la stessa del modello *V-to-C* appena descritto, salvo distaccarsene coll'imporre a *V2ons* di sincronizzare non più con *Ccent*, ma con *Crel*. L'effetto sostanziale è quello di spostare la durata del ciclo di V_2 a destra, sottraendolo meglio alla sovrapposizione consonantica.

4.4. Analisi statistica

Le analisi statistiche descrittive sono state condotte mediante il software *Jamovi* (<https://www.jamovi.org/>, versione 2.2.5), importando le tabelle direttamente da *Excel*. L'analisi statistica inferenziale è stata fatta precedere dal test di normalità “*Shapiro-Wilk*”, che ha rivelato come alcune misure non fossero distribuite in modo normale. È stato dunque utilizzato il test non parametrico *Mann-Whitney*. La variabile dipendente considerata è stata sempre la “geminazione” (GEMINATION), mentre

le variabili dipendenti, di tipo continuo, sono rappresentate dalle misure temporali (puntuali o sotto forma di intervalli). Inoltre, sono stati utilizzati i filtri di *Jamovi*, per produrre statistiche divise per soggetto (AG, MP, FC, BG).

5. Risultati

5.1 Ruolo degli intervalli consonantici nella geminazione

In Tab. 1 si riportano le medie dei singoli intervalli consonantici indagati per comprendere il loro ruolo nella geminazione, ossia *Cons_Ctarg*, *Ctarg_Crel* e *Crel_Creloff*. Per tutti e quattro i soggetti, la totalità degli intervalli mostra una differenza significativa tra consonante geminata e scempia alla velocità d'elocuzione "normale".

Tabella 1 - *Media intervalli consonantici principali nei singoli soggetti. Unità di misura ms; G=geminates; S= singletons. *=coppie $p < .05$. Test Mann-Whitney U*

	AG		MP		FC		BG	
<i>Time Intervals</i>	G	S	G	S	G	S	G	S
<i>Cons_Ctarg</i>	111*	89*	90*	80*	102*	89*	122*	91*
<i>Ctarg_Crel</i>	42*	17*	51*	17*	45*	15*	60*	14*
<i>Crel_Creloff</i>	92*	68*	87*	74*	101*	76*	99*	78*

Questi dati smentiscono l'ipotesi che la maggior durata delle geminate rispetto alle scempie dell'italiano sia dovuta solo all'aumento della fase di tenuta dell'occlusione.

5.2 Analisi dei dati sperimentali "reali" per gli intervalli "critici"

Le medie degli intervalli da noi scelti perché in grado di rivelare, come una cartina al tornasole, le conseguenze diverse delle sincronizzazioni previste dai diversi modelli, suddivise per soggetto sono sintetizzati nella Tab. 2, che evidenzia anche se la differenza tra scempie e geminate all'interno di ciascun soggetto è significativa ($p < .05$; mediante asterisco).

Tabella 2 - *Media dei soggetti per gli intervalli temporali (arrotondati a ms; G = geminates; S = Singletons. *=la differenza nelle coppie è significativa per $p < .05$. Test Mann-Whitney U*

	AG		MP		FC		BG	
<i>Time Intervals</i>	G	S	G	S	G	S	G	S
<i>V1targ_Cons</i>	3	21	47*	65*	70*	90*	45*	81*
<i>V1targ_Ccent</i>	137*	115*	166	153	193	188	183	179
<i>Ccent_V2rel</i>	130	139	110*	92*	141*	119*	153*	96*
<i>Creloff_V2rel</i>	65	64	8	10*	15*	38	13	11
<i>V1rel_V2targ</i>	210*	159*	167*	119*	208*	151*	207*	132*

I risultati in questa tabella dipingono una situazione in cui, per la maggioranza dei soggetti (almeno tre su quattro), l'accorciamento della prima vocale è ottenuto tramite l'anticipazione dell'inizio consonantico (*V1targ_Cons*) ma senza anticipare il baricentro della tenuta dell'occlusione (*V1targ_Ccent*). Il mantenimento della durata della seconda vocale è realizzato tramite una posticipazione della fine della seconda vocale (*Ccent_V2rel*) e da un mantenimento dell'intervallo tra la fine della consonante e la fine della seconda vocale (*Creloff_V2rel*). Si assiste inoltre a un allungamento della durata della transizione tra la fine della prima vocale e il raggiungimento del target articolatorio della seconda (*V1rel_V2targ*).

5.3 Confronto con i modelli cinematici a partire dai dati reali

La Tab. 3 mostra se le predizioni fatte da ciascun modello per ogni intervallo al variare della lunghezza consonantica (=: non varia; ≠: varia) si conformano o meno ai risultati ottenuti (=: non varia; ≠: varia).

L'intervallo *V1targ_Cons* si mostra differente tra VCV e VCCV in 3 soggetti su 4, come predetto da tutti i modelli.

V1targ_Ccent resta invariato in 3 soggetti su 4, diversamente dalle predizioni dei modelli *VtoC* e *Hybrid*. L'intervallo *Ccent_V2rel*, invece, è risultato significativamente diverso in 3 soggetti su 4, anche qui divergendo dalle predizioni del modello *VtoC*.

Creloff_V2rel si è rivelato simile tra VCV e VCCV in 3 soggetti su 4, in conformità alle predizioni del modello *VtoC* e *Hybrid*. Infine, *V1rel_V2targ* si è mostrato significativamente diverso tra scempia e geminata in tutti i soggetti, conformemente alle predizioni dei modelli *VtoC* e *Hybrid*.

Tabella 3 - Risposta dei soggetti alle predizioni dei modelli *VtoV* or *VtoC* (=: VCV non significativamente diverso da VCCV; ≠: VCV significativamente diverso da VCCV; $p < .05$)

	<i>VtoV</i>	<i>VtoVmod</i>	<i>VtoC</i>	<i>Hybrid</i>	<i>AG</i>	<i>MP</i>	<i>FC</i>	<i>BG</i>
<i>V1targ_Cons</i>	≠	≠	≠	≠	=	≠	≠	≠
<i>V1targ_Ccent</i>	=	=	≠	≠	≠	=	=	=
<i>Ccent_V2rel</i>	≠	≠	=	≠	=	≠	≠	≠
<i>Creloff_V2rel</i>	≠	≠	=	=	=	=	≠	=
<i>V1rel_V2targ</i>	=	=	≠	≠	≠	≠	≠	≠

5.4 Confronto tra i dati ottenuti dalla "Geminata Reale (GR)" e i modelli cinematici simulati con la "Geminata Virtuale (GV)"

Nelle tabelle che seguono, mostreremo i risultati del confronto tra i dati ricavati dalle simulazioni delle conseguenze sull'organizzazione temporale predette da ciascuno dei quattro modelli di organizzazione temporale della GV e i dati ottenuti per la geminata reale, per ciascuno dei 5 intervalli critici. Ovviamente, la logica sortesa a questo confronto è che se la simulazione relativa a un certo modello non produce dati temporali significativamente diversi per l'intervallo in questione rispetto ai dati realmente ottenuti, allora il modello risulta confermato.

La Tab. 4 mostra i valori temporali (ms) ottenuti con la simulazione del modello *V-to-V* (GV) affiancati ai dati misurati dalla produzione reale (GR), soggetto per soggetto per ciascuno dei 5 intervalli.

Tabella 4 - *Medie dei soggetti per gli intervalli temporali critici delle geminate "reali" (GR) e virtuali (GV), simulate in base al modello V-to-V (arrotondati a ms; G = geminates; S = Singletons. *=la differenza nelle coppie è significativa per $p < .05$. Test Mann-Whitney U*

<i>VtoV</i>	<i>AG</i>		<i>MP</i>		<i>FC</i>		<i>BG</i>	
<i>Time Intervals</i>	<i>GR</i>	<i>GV</i>	<i>GR</i>	<i>GV</i>	<i>GR</i>	<i>GV</i>	<i>GR</i>	<i>GV</i>
<i>V1targ_Cons</i>	3	8	47*	56*	70*	83*	45*	74*
<i>V1targ_Ccent</i>	137*	113*	166	154	193	187	183	179
<i>Ccent_V2rel</i>	130	140	110*	92*	141	121	153*	96*
<i>Creloff_V2rel</i>	65	55	8	1	15	30	13	4
<i>V1rel_V2targ</i>	210*	159*	167*	119*	208*	151*	207*	132*

La Tab. 5 esibisce i valori temporali (ms) ottenuti con la simulazione del modello *V-to-Vmod*, che è stato modificato tramite l'allineamento forzato del centro della tenuta articolatoria (*Ccent*) con l'inizio del movimento verso la seconda vocale (*V2ons*). I dati simulati (GV) sono affiancati ai dati misurati dalla produzione reale (GV), soggetto per soggetto per ciascuno dei 5 intervalli.

Tabella 5 - *Medie dei soggetti per gli intervalli temporali critici delle geminate "reali" (GR) e virtuali (GV), simulate in base al modello V-to-V modificato (arrotondati a ms; G = geminates; S = Singletons. *=la differenza nelle coppie è significativa per $p < .05$. Test Mann-Whitney U*

<i>VtoVmod</i>	<i>AG</i>		<i>MP</i>		<i>FC</i>		<i>BG</i>	
<i>Time Intervals</i>	<i>GR</i>	<i>GV</i>	<i>GR</i>	<i>GV</i>	<i>GR</i>	<i>GV</i>	<i>GR</i>	<i>GV</i>
<i>V1targ_Cons</i>	3*	-41*	47*	19*	70*	38*	45	29
<i>V1targ_Ccent</i>	137*	69*	166*	116*	193*	142*	183*	134*
<i>Ccent_V2rel</i>	130*	188*	110*	130*	141	166	153*	141*
<i>Creloff_V2rel</i>	65*	104*	8*	39*	15*	76*	13*	49*
<i>V1rel_V2targ</i>	210*	159*	167*	119*	208*	151*	207*	132*

La Tab. 6 mostra i valori temporali (ms) ottenuti con la simulazione del modello *V-to-C*. I dati simulati (GV) sono affiancati ai dati misurati dalla produzione reale (GR), soggetto per soggetto per ciascuno dei 5 intervalli.

Tabella 6 - Medie dei soggetti per gli intervalli temporali critici delle geminate "reali" (GR) e virtuali (GV), simulate in base al modello V-to-C (arrotondati a ms; G = geminates; S = Singletons. * = la differenza nelle coppie è significativa per $p < .05$. Test Mann-Whitney U

<i>VtoC</i>	<i>AG</i>		<i>MP</i>		<i>FC</i>		<i>BG</i>	
<i>Time Intervals</i>	<i>GR</i>	<i>GV</i>	<i>GR</i>	<i>GV</i>	<i>GR</i>	<i>GV</i>	<i>GR</i>	<i>GV</i>
<i>V1targ_Cons</i>	3 *	-23 *	47 *	36 *	70	53	45	43
<i>V1targ_Ccent</i>	137 *	83 *	166 *	133 *	193 *	157 *	183 *	148 *
<i>Ccent_V2rel</i>	130 *	188 *	110 *	130 *	141	166	153 *	141 *
<i>Creloff_V2rel</i>	65 *	104 *	8 *	39 *	15 *	76 *	13 *	49 *
<i>V1rel_V2targ</i>	210 *	177 *	167 *	136 *	208 *	165 *	207 *	146 *

La Tab. 7 esibisce i valori temporali (ms) ottenuti con la simulazione del modello "ibrido". I dati simulati (GV) sono affiancati ai dati misurati dalla produzione reale (GR), soggetto per soggetto per ciascuno dei 5 intervalli.

Tabella 7 - Medie dei soggetti per gli intervalli temporali critici delle geminate "reali" (GR) e virtuali (GV), simulate in base al modello Hybrid (arrotondati a ms; G = geminates; S = Singletons. * = la differenza nelle coppie è significativa per $p < .05$. Test Mann-Whitney U

<i>Hybrid</i>	<i>AG</i>		<i>MP</i>		<i>FC</i>		<i>BG</i>	
<i>Time Intervals</i>	<i>GR</i>	<i>GV</i>	<i>GR</i>	<i>GV</i>	<i>GR</i>	<i>GV</i>	<i>GR</i>	<i>GV</i>
<i>V1targ_Cons</i>	3 *	-23 *	47 *	36 *	70	53	45	43
<i>V1targ_Ccent</i>	137 *	83 *	166 *	133 *	193 *	157 *	183 *	148 *
<i>Ccent_V2rel</i>	130 *	205 *	110 *	147 *	141 *	181 *	153	155
<i>Creloff_V2rel</i>	65 *	120 *	8 *	56 *	15 *	90 *	13 *	63 *
<i>V1rel_V2targ</i>	210	193	167 *	153 *	208	180	207 *	160 *

La Tab. 8 indica per quali soggetti la simulazione in base a ciascun modello ha prodotto un valore temporale non significativamente diverso dai valori reali.

Tabella 8 - Soggetti la cui geminata virtuale non ha una durata statisticamente diversa dalla geminata reale

	<i>VtoV</i>	<i>VtoVmod</i>	<i>VtoC</i>	<i>Hybrid</i>
<i>V1targ_Cons</i>	AG	BG	FC,BG	FC,BG
<i>V1targ_Ccent</i>	MP,FC,BG	-	-	-
<i>Ccent_V2rel</i>	AG,FC	FC	FC	BG
<i>Creloff_V2rel</i>	AG,MP,FC,BG	-	-	-
<i>V1rel_V2targ</i>	-	-	-	AG,FC

Come è facilmente osservabile, la simulazione di maggior successo è quella relativa al modello V-to-V, in base al quale un semplice allungamento simmetrico a destra e

a sinistra del *Ccenter* della consonante scempia basta a rendere conto dell'intervallo tra il raggiungimento del target articolatorio della prima vocale e il centro della fase di tenuta articolatoria, di quello tra quest'ultimo e il rilascio della consonante (almeno per due soggetti), e, per tutti i soggetti, dell'intervallo relativo allo "spazio vitale" della seconda vocale, liberata dalla sovrapposizione con la consonante. L'unico intervallo non riprodotto è quello relativo alla durata della transizione tra la posizione articolatoria della prima vocale e quella della seconda, che per il modello *V-to-V* deve rimanere necessariamente invariato tra VCV e VCCV (la variazione interessando il solo ciclo consonantico), ma questo era un dato atteso dalle rilevazioni reali.

Un'altra osservazione interessante di tipo generale è l'estrema idiosincronicità delle risposte. All'interno di ciascun modello, salvo il *V-to-V*, i soggetti che producono una simulazione di successo per un intervallo non sono mai gli stessi per gli altri intervalli (con l'eccezione, forse di FC).

Da notare anche il fallimento della simulazione con il modello *V-to-C*.

6. *Discussione e conclusioni*

È nostra opinione che i dati che abbiamo presentato gettino una nuova luce sulla natura delle consonanti geminate italiane e propongano un metodo originale per lo studio dell'organizzazione temporale intra- e inter-gestuale come quello della geminata virtuale. Ricordiamo brevemente che l'impostazione di questo studio si è basata sull'individuazione di alcuni intervalli temporali che mettono in relazione dei *landmark*, nella terminologia di Gafos (2002), che non sono altro che punti di ancoraggio nel ciclo articolatorio per la sincronizzazione temporale di gesti consonantici e vocalici di tipo sia intra-sillabico (come la relazione C-V) che inter-sillabico (come la relazione V-C nel bisillabo VCV), sia solo consonantici (C-C in VCCV) che solo vocalici (V-V in V(C)CV). Gli intervalli individuati costituiscono dei veri banchi di prova del funzionamento di alcuni modelli di organizzazione temporale intra-e inter-sillabica presenti in letteratura. I modelli analizzati sono il modello *V-to-V* originalmente proposto da Öhman (1967), lo stesso modello modificato tramite una sincronizzazione del centro del ciclo consonantico "in fase" con l'inizio della seconda vocale (*V-to-Vmod*), il modello *V-to-C* di Browman, Goldstein (1989, 2000) e un modello ibrido da noi proposto sulla base dei risultati ottenuti da Zmarich et al. (2007) e Gili Fivela et al. (2007).

Riassumendo, ci si è posti tre obiettivi: testare i modelli *V-to-V* e *V-to-C*, basandosi su intervalli inerenti a dati reali; verificare il ruolo dei singoli intervalli consonantici nel contrasto di lunghezza consonantica; verificare il funzionamento dei modelli di temporizzazione *V-to-V* e *V-to-C* mediante la costruzione di una geminata virtuale.

L'analisi dei dati cinematici "reali" prodotti dai soggetti nell'esperimento di Zmarich et al. (2011) ha decretato che, per la maggioranza dei soggetti (almeno tre su quattro), l'accorciamento della prima vocale è ottenuto tramite l'anticipazione dell'inizio consonantico (*V1targ_Cons*) ma senza anticipare il baricentro della

tenuta dell'occlusione ($V1_{targ_Ccent}$). Il mantenimento della durata della seconda vocale è realizzato tramite una posticipazione della fine della seconda vocale ($Ccent_V2_{rel}$) e da un mantenimento dell'intervallo tra la fine della consonante e la fine della seconda vocale ($C_{rel}_{off_V2_{rel}}$). Si assiste, inoltre, a un allungamento della durata della transizione tra la fine della prima vocale e il raggiungimento del target articolatorio della seconda ($V1_{rel_V2_{targ}}$). Vale la pena notare che l'allungamento di questo intervallo è l'unica variazione tra singola e geminata che è condivisa da tutti i soggetti.

Un risultato interessante, e mai presentato prima, riguarda la verifica dell'intervallo di transizione vocalica $V1_{rel_V2_{targ}}$: il risultato per quest'ultimo intervallo, mai indagato nei precedenti contributi di Zmarich, Gili Fileva e colleghi, confuta l'ipotesi avanzata da Smith (1995), secondo la quale la geminazione italiana si conforma senza riserve al modello V -to- V di Öhman (1967). Infatti, diversamente da quanto predetto dalla studiosa, la fase di transizione vocalica del ciclo vocalico $V1_{rel_V2_{targ}}$ si è mostrata significativamente differente per tutti i soggetti, da quella presente nella scempia. Il risultato sostanzia e spiega quanto rilevato da Zmarich et al. (2011) per la maggior durata del ciclo vocalico $V1_{targ_V2_{rel}}$ ($V1_{ons_V2_{off}}$, nella terminologia di Zmarich e colleghi), nel senso che questo allungamento, non previsto dal modello V -to- V , è dovuto principalmente all'aumento della durata della transizione vocalica (anche questo non previsto dal modello).

Circa il secondo obiettivo, si è evidenziato come nella geminazione sia coinvolto un aumento di durata di ciascun singolo intervallo consonantico ($Cons_C_{targ}$, $C_{targ_C_{rel}}$ e $C_{rel_C_{rel}_{off}}$). In primo luogo, tale dato confuta l'idea comune che la geminazione sia causata da un mero allungamento della sola fase di tenuta articolatoria ($C_{targ_C_{rel}}$). Inoltre, sebbene lo studio si focalizzi puramente sulle durate temporali, questo risultato potrebbe portare a ipotizzare che il mutamento da geminata a scempia, in italiano, possa essere spiegato dal cosiddetto *change by shrinking* (Cho, 2006), oppure dal cosiddetto *change by stiffness* (Cho, 2006). In altri termini, nel primo caso la scempia, rispetto alla geminata, presenterebbe una modifica di tutti i parametri dell'equazione massa-molla (*stiffness* e ampiezza), che causerebbero una maggiore durata della geminata (ricordiamo che nell'equazione massa-molla la durata non costituisce un parametro (da pianificare) ma è un (sotto) prodotto implicito dello svolgimento del processo descritto dall'equazione). Nel secondo caso, la geminata, rispetto alla scempia, presenterebbe una modifica della sola *stiffness*. Per avvalorare tale predizione, tuttavia, sarebbe auspicabile effettuare nuove misurazioni, volte a determinare il ruolo delle altre variabili cinematiche come l'ampiezza, la velocità massima *time-to-peak velocity* (che indica la proporzione della fase di accelerazione su quella di decelerazione), che sono attualmente in fase di studio.

Una parte importante e metodologicamente nuova di questo studio ha riguardato la simulazione delle organizzazioni temporali dei gesti cinematici per consonanti e vocali nel bisillabo VCCV in base ai quattro modelli e il loro confronto con i dati delle geminate "reali". Anche qui siamo partiti dalla formalizzazione

di Gafos (2002) che modella la geminata in *close transition*, come quella italiana, raddoppiando la curva della scempia ricavata dai dati reali ottenuti da Zmarich et al. (2011) e sovrapponendola ai profili cinematici del ciclo vocalico V-V dello stimolo *singleton*. Il confronto ha messo in luce che, nonostante i problemi del modello *V-to-V* esposti in precedenza, quest'ultimo ha approssimato con più successo il tipo di sincronizzazione che compare nelle geminate "reali".

In conclusione, non va nascosto che la consonante labiale qui indagata presenta un ridotto grado di complessità poiché la consonante e le vocali contestuali sono realizzate da articolatori diversi (rispettivamente, labbra e dorso della lingua) i cui movimenti sono reciprocamente indipendenti, e che altri tipi di consonanti linguali aumenterebbero il grado di complessità e di non linearità della sincronizzazione tra consonanti e vocali, ma proprio per questo motivo l'indipendenza articolatoria garantisce un ambiente ideale per lo studio di un tipo di sincronizzazione "pura".

Bibliografia

- BERTINETTO, P.M. (1981). Strutture Prosodiche dell'Italiano, Studi di grammatica italiana, *Accademia della Crusca*, Firenze.
- BROWMAN, C., GOLDSTEIN, L. (1988). Some notes on syllable structure in Articulatory Phonology. In *Phonetica*, 45(2-4), 140-155.
- BROWMAN, C., GOLDSTEIN, L. (1989). Articulatory gestures as phonological units. *Phonology*, 6, 201-251.
- BROWMAN, C., GOLDSTEIN, L. (1990). Tiers in Articulatory Phonology, with some implications for casual speech. In KINGSTON, J., BECKMAN, M.E. (Eds.), *Papers in Laboratory Phonology I: Between the grammar and physics of speech*. Cambridge: Cambridge University Press, 340-376.
- BROWMAN, C., GOLDSTEIN, L. (1992). Articulatory Phonology: An Overview. In *Phonetica*, 49, 155-180.
- BROWMAN, C., GOLDSTEIN, L. (2000). Competing constraints on intergestural coordination and self-organization of phonological structures. In *Bulletin de la Communication Parlée*, 5, 25-34.
- BURRONI, F., MASPOG, S., BENKER, N., HOOLE, P. & KIRBY, J. (2024). Spatiotemporal Features of Bilabial Geminate and Singleton Consonants in Italian. *Proceedings of the 13th International Seminar of Speech Production, 13-17 May 2024 Autrans FR*. 189-192.
- CELATA, C., MELUZZI, C. & BERTINI, C. (2022). Acoustic and Kinematic Correlates of Heterosyllabicity in Different Phonological Contexts. In *Language and Speech*, 65(3), 755-780.
- CHO, T. (2002). *The Effects of Prosody on Articulation in English*. New York, London: Routledge.
- DI BENEDETTO, M.G., DE NARDIS, L. (2021A). "Gemination in Italian: The affricate and fricative case," *Speech Commun.* (in press); arXiv:2005.06959.
- DIPINO, D., CELATA, C. (2018). An UTI study of alveolar stops in Italian. In VIETTI, A., SPREAFICO, L., MEREU, D. & GALATÀ, V. (Eds.), *Il parlato nel contesto naturale. Speech in the natural context*. Milano: Officinaventuno, 41-53.

- ESPOSITO, A., DI BENEDETTO, M.G. (1999). Acoustical and perceptual study of gemination in Italian stops. In *Journal of The Acoustical Society of America*, 106, 2051-2062.
- FOWLER, C. (1980). Coarticulation and theories of extrinsic timing. In *Journal of Phonetics*, 8, 113-133.
- GAFOS, A. (2002). A grammar of gestural coordination. In *Natural Language and Linguistic Theory*, 20(2), 269-337.
- GAFOS, A., GOLDSTEIN, L. (2012). Articulatory representation and organization. In COHN, A., FOURGERON, C. & HUFFMAN, E.M. (Eds.), *Oxford Handbook of Laboratory Phonology*. Oxford: Oxford University Press, 220-231.
- 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.
- HAGEDORN, C., PROCTOR, M. & GOLDSTEIN, L. (2011). Automatic Analysis of Singleton and Geminate Consonant Articulation Using Real-time Magnetic Resonance Imaging. In *Journal of the Acoustical Society of America*, 130(4), EL251-EL257.
- HALL, N. (2010). Articulatory Phonology. In *Language and Linguistics*, 4(9), 818-830.
- KRAUSE, P.A., KAWAMOTO, A.H. (2020). On the timing and coordination of articulatory movements: Historical perspectives and current theoretical challenges. In *Language and Linguistics*, 14(6), e12373.
- LAHIRI, A., HANKAMER, J. (1988). The timing of geminate consonants. In *Journal of Phonetics*, 16, 327-338.
- LOPORCARO, M. (1990). On the analysis of geminates in Standard Italian and Italian dialects. In HURCH, B., RHODES, R.A. (Eds.), *Natural Phonology. The State of the Art*. Berlin, New York: Mouton de Gruyter, 153-188.
- LÖFQVIST, A. (2005). Lip kinematics in long and short stop and fricative consonants. In *Journal of the Acoustical Society of America*, 117, 858-878.
- MAIRANO P., DE IACOVO, V. (2019). Gemination in Northern versus Central and Southern Varieties of Italian: A Corpus-based Investigation. In *Language and Speech*, 63(3): 608-634.
- MARIN, S. (2013). Organization of Complex Onsets and Codas in Romanian: A gestural approach. In *Journal of Phonetics*, 41, 211-227.
- MAROTTA, G., VANELLI, L. (2021). *Fonologia e prosodia dell'italiano*. Roma: Carocci.
- MATTEI, M., DI BENEDETTO, M.G. (2000). Acoustic analysis of singleton and geminate nasals in Italian. *Web-SLS The European Journal of Language and Speech*, 1-11. Retrieved from <http://wrangler.essex.ac.uk/web-sls>
- PAYNE, E.M. (2005). Phonetic variation in consonant gemination. In *Journal of the International Phonetic Association*, 35, 153-189.
- PAYNE, E.M. (2006). Non-durational indices in Italian geminate consonants. In *Journal of the International Phonetic Association*, 36(1), 83-95.

- PASTÄTTER, M., POUPLIER, M. (2015). Onset-vowel timing as a function of coarticulation resistance: Evidence from articulatory data. *Proceedings of the 18th International Congress of Phonetic Sciences, Glasgow*.
- POUPLIER, M. (2020). Articulatory phonology. In *Oxford research encyclopedia of linguistics*. <https://doi.org/10.1093/acrefore/9780199384655.013.745>
- SALTZMAN, E.L. (1986). Task dynamic coordination of the speech articulators: a preliminary model. Generation and modulation of action patterns, *Experimental Brain Research*, Series 15, 129-144.
- SMITH, C.L. (1995). Prosodic patterns in the coordination of vowel and consonant gestures. In CONNELL, B., ARVANITI, A. (Eds.), *Phonology and Phonetic Evidence. Papers in Laboratory Phonology IV*. Cambridge: Cambridge University Press, 205-222.
- STEVENS, M., HAJEK, J. (2005). Raddoppiamento sintattico (RS) and word-medial gemination in Italian. In *Theoretical and Experimental Approaches to Romance Linguistics: Selected Papers from the 34th Linguistic Symposium on Romance Languages (LSRL)*, Salt Lake City, March 2004 (Vol. 272, p. 257-271). Amsterdam, Netherlands: John Benjamins.
- TURVEY, M.T. (1990). Coordination. In *American Psychologist*, 45, 938-953.
- ZMARICH, C., GILI FIVELA, B. (2005). Consonanti scempie e geminate in italiano. Studio cinematico e percettivo dell'articolazione bilabiale e labiodentale. In COSÌ, P. (Ed.), *La misura dei parametri. Aspetti tecnologici ed implicazioni nei modelli linguistici. Atti del I Convegno Nazionale AISV – Associazione Italiana di Scienze della Voce*, Padova, 2-4 dicembre 2004. Torriana (RN), EDK Editore, 429-448.
- 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 GIORDANI, V., BRUSEGHINI, V. & COSÌ, P. (a cura di), *Atti del III Convegno Nazionale dell'Associazione Italiana di Scienze della Voce (AISV)*, Trento, 29-30/11-1/12/2006, EDK Editore srl, Torriana (RN), 151-163, 2006.
- 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 GIORDANI, V., BRUSEGHINI, V. & COSÌ, P. (Eds.), *Scienze Vocali e del linguaggio – Metodologie di valutazione e risorse linguistiche. Atti del III Convegno Nazionale AISV – Associazione Italiana di Scienze della Voce, Povo (Trento), 29 novembre – 1 dicembre 2006*. Torriana (RN), EDK Editore, 151-163.
- ZMARICH, C., GILI FIVELA, B., PERRIER, P., SAVARIAUX, C. & TISATO, G. (2011). Speech Timing Organization for the Phonological Length Contrast in Italian Consonants. In COSÌ, P., DE MORI, R., DI FABBRIZIO, G. & PIERACCINI, R. (Eds.), *Proceedings of Interspeech 2011: 12th Annual Conference of the International Speech Communication Association*, Florence, Italy, August 27-31, 2011. 401-404.

