DOI: 10.17469/O2102AISV000004

SONIA D'APOLITO, BARBARA GILI FIVELA

Targetless schwa in francese L2: primi risultati in area italofona

This paper focuses on how non-native consonant clusters are produced by Italian learners of French L2 depending on speech rate, prosodic conditions and sequence type. We compare the productions by three Italian learners and one native speaker, paying particular attention to epenthetic vowels, which are possibly inserted between the two consonants. We examine the acoustic and articulatory characteristics of these contexts to observe how consonant gestures are coordinated and if the insertion of a vowel corresponds to a full vowel insertion (i.e. there is an articulatory target) or to a gestural mistiming (i.e. there is no articulatory specific vowel target). Results show that the inserted vowel has a high variability in acoustic quality and duration; moreover, on the articulatory level, no target has been identified for the tongue dorsum gesture and different patterns are actually realized because of the influence of speech rate, prosodic conditions and sequence type. Our results suggest that vowel insertion in Italian speakers production seems to be due to the failure in reaching a correct coordination between the two consonantal gestures (gestural mistiming).

Key words: inserzione vocalica, vocale epentetica e intrusiva, coordinazione intragestuale, dati articolatori EMA, francese L2.

Introduzione

L'inserzione vocalica può avvenire in base a due meccanismi articolatori: i) inserzione di un gesto vocalico articolatorio – vocale epentetica; ii) un gesto di transizione tra due gesti consonantici che produce un suono vocalico – vocale intrusiva. Secondo Hall (2003; 2006; 2011), le vocali intrusive mostrano una qualità variabile, poiché possono avere le caratteristiche di una vocale schwa, oppure essere più simili alla vocale adiacente o essere influenzate dal luogo di articolazione delle consonanti adiacenti. Peraltro, Levin (1987) afferma che una vocale intrusiva può anche esibire le qualità di una vocale non presente nel sistema linguistico, quindi, una vocale schwa può essere realizzata tra due consonanti benché non appartenga al sistema linguistico in esame. Inoltre, la durata della vocale è variabile ed è fortemente influenzata dalla velocità di eloquio, al punto che a velocità sostenuta una vocale intrusiva non si realizza. La sua funzione è quella di favorire la percezione dei segmenti di un nesso consonantico. Al contrario, una vocale epentetica ha una qualità fissa o caratteristiche simili alla vocale adiacente, il suo inserimento non è influenzato dalla velocità di eloquio e ha la funzione è di riparare una struttura marcata.

Per quanto riguarda la durata acustica, una vocale epentetica ha una durata più lunga di una vocale intrusiva. Ad esempio, nella lingua Salish si realizzano entrambe le vocali,

ossia degli schwa epentetici, fonologicamente presenti sia dal punto di vista acustico che percettivo, e degli schwa intrusivi che si presentano come elementi di transizione. Uno studio acustico ha dimostrato che le vocali epentetiche hanno effettivamente una durata maggiore rispetto a quelle intrusive (Shanin, Kimary & Black, 2004). Visto che la durata di una vocale epentetica è maggiore, ci si aspetta quindi che anche la durata di una sequenza sia maggiore in caso di inserimento di una vocale epentetica. Davidson, Roon (2008) in effetti hanno osservato nessi consonantici in posizione iniziale nella lingua russa e i loro risultati confermano che sequenze #CəC, con inserimento di uno schwa fonologico, hanno una durata maggiore rispetto alle sequenze con uno schwa intrusivo. La durata delle sequenze è stata anche osservata confrontando lo schwa intrusivo con schwa lessicalizzati (con un target sottostante), in produzioni da parte di parlanti americani di sequenze non native della lingua polacca (es./sC-, səC- zC-/; Davidson, 2005) e della lingua ceca (es. /C_1əC_2/, in cui C_1 è una fricativa combinata con tutte le possibili consonanti; Davidson, 2006). I risultati mostrano che le sequenze con uno schwa intrusivo hanno una durata minore rispetto alle sequenze con schwa lessicalizzato. Inoltre, lo schwa intrusivo presenta valori di F1 e F2 più bassi rispetto allo schwa lessicalizzato, per via della coarticolazione con la vocale successiva. Davidson interpreta questi risultati affermando che uno schwa intrusivo si realizza attraverso un *gestural mistiming*, cioè i parlanti falliscono nel coordinare in modo appropriato, di fatto in termini di sovrapposizione, i gesti consonantici della sequenza. A supporto di questa interpretazione ci sono anche gli studi con dati ricavati mediante *Ultrasound Tongue Imaging* (Davidson, 2005; Davidson, Stone, 2003), in cui è stata osservata la conformazione della lingua durante la produzione di [zᵊC], con schwa intrusivo, di [səC], con schwa lessicalizzato e di [sC]. I risultati mostrano che la configurazione della lingua durante la produzione di [zᵊC] è più simile a [sC] rispetto a [səC] poiché [s] seguita da [ə] ha una posizione di partenza più bassa nel cavo orale essendo coarticolata con la vocale [ə] che ha una posizione del corpo della lingua più bassa. Al contrario, nelle sequenze [sC] e [zᵊC], la fricativa alveolare si coarticola direttamente con la consonante successiva. Ad esempio, [s] seguita da [k] assume una posizione più alta poiché il dorso della lingua durante [k] è innalzato, sia nella sequenza [sk] che [zᵊk] con schwa intrusivo (Davidson, 2005). I risultati dimostrano che i parlanti non necessariamente riparano una sequenza marcata con uno schwa epentetico, ma falliscono nel coordinare i gesti consonantici, per cui tra le due consonanti si crea uno spazio in cui si realizza uno schwa di transizione. I risultati dello studio precedente, indicano inoltre che uno schwa intrusivo non è specificato da un proprio target articolatorio dal momento che le due consonanti si coarticolano direttamente, al contrario di quanto accade nel caso in cui lo schwa sia una vocale vera e propria. Infatti, come abbiamo detto precedentemente, una vocale epentetica si realizza con un gesto articolatorio con *target* vocalico, mentre una vocale intrusiva si realizza grazie a un gesto di transizione tra due gesti consonantici. Infine, Browman e Goldstein (1992) hanno osservato il movimento della lingua durante la realizzazione di una vocale schwa in posizione mediana all'interno della pseudo-parola [pipəpapə] in americano. La posizione della lingua durante la produzione dello schwa mediano non è una semplice interpolazione tra consonanti e vocali adiacenti. Questo fa pensare che

in inglese non si possa escludere la presenza di un target per lo schwa. In altre lingue, invece, come l'arabo marocchino, la vocale schwa sembra derivare da gesti consonantici non sufficientemente sovrapposti (Gafos, 2002) e non dalla presenza di un bersaglio articolatorio di tipo vocalico. Per quanto riguarda la lingua francese, Rialland (1986) distingue due tipi di schwa individuati in: 1) un morfema interno che ha un nucleo fonologico e 2) un morfema finale che deriva dal rilascio acustico della consonante. Non ci risulta però che siano disponibili dati articolatori relativi alla presenza o meno di bersagli per le vocali.

In base alla Fonologia articolatoria (Browman, Goldstein, 1986; 1987; 1989; 1992; 2007; Browman, 1995), la presenza di un *target* per la vocale e la scarsa sovrapposizione di gesti consonantici corrispondono alla presenza, rispettivamente, di vocali specificate e non specificate: una vocale non specificata non ha *target* articolatorio (*targetless*), mentre una vocale specificata ha un proprio bersaglio. Pertanto, nella produzione di sequenze consonantiche non-native potrebbe realizzarsi questo tipo di meccanismo, cioè i parlanti, per riparare un nesso marcato (Eckman, 2008), potrebbero anche realizzare un tipo di coordinazione caratterizzato da un basso livello di sovrapposizione gestuale tra i gesti consonantici che fa sì che si crei uno spazio vocalico aperto sufficiente per l'inserzione vocalica (Davidson, 2005).

1. *Obiettivi e Ipotesi*

L'obiettivo di questo studio acustico-articolatorio è quello di osservare le sequenze di sibilanti (alveolare-postalveolare - AP - e postalveolare-alveolare - PA) a confine di parola, realizzate mediante l'inserimento di una vocale d'appoggio da parte di apprendenti italiani di francese L2 avanzato. In italiano, le sequenze di consonanti al confine di parola compaiono in pochi casi, principalmente in preposizioni, prestiti stranieri, o nei verbi in la vocale finale sia stata troncata (Muliačić, 1973; Farnetani, Busà, 2004). In francese, invece, le sequenze consonantiche a confine di parola sono molto frequenti e quelle di sibilanti possono anche essere dominio di assimilazioni di luogo (Niebuhr, Lancia & Meunier, 2008). Di conseguenza, gli apprendenti avranno difficoltà nel realizzare queste sequenze e una delle possibili strategie per riparare i nessi potrà essere quella di inserire una vocale d'appoggio. Peraltro, in molte varietà di francese, tra cui il francese del Sud della Francia, è previsto uno schwa e nella didattica del francese in Italia molti insegnanti producono questo schwa. Di conseguenza, la maggior parte degli italofoni apprendenti francese è convinta che lo schwa sia standard. Del resto alcuni materiali didattici trascrivono tuttora parole (es. *tasse*) con uno schwa finale. In particolare, in questo studio si vuole osservare dal punto di vista articolatorio l'*output* acustico corrispondente a una eventuale vocale d'appoggio per capire se esso sia dovuto all'inserimento di una vocale specificata da un proprio *target* articolatorio – vocale epentetica con *target* – oppure sia *targetless* – vocale intrusiva dovuta a un *gestural mistiming*. L'ipotesi è che l'inserzione vocalica sia dovuta ad un'epentesi come strategia generale adottata dagli apprendenti in base alle caratteristiche fonetiche e fonologiche della L1. Tuttavia, dal punto di vista articolatorio, non si esclude un processo di *gestural*

mistiming dovuto alla difficoltà da parte degli apprendenti di coordinare in modo appropriato i due gesti consonantici. Inoltre, la produzione delle sequenze sarà osservata al variare della struttura prosodica e della velocità di eloquio. Circa l'influenza della condizione prosodica, la presenza di un confine prosodico interferisce con il processo coarticolatorio (Byrd, Choi, 2006), mentre una velocità di eloquio sostenuta facilita la coproduzione di segmenti successivi (Byrd, Tan, 1996). Ci si aspetta, quindi, che l'inserzione vocalica, intrusiva e epentetica, possa essere favorita in caso di confine intonativo mentre potrebbe ridursi, se non addirittura scomparire, a velocità sostenuta. Infine, circa il tipo di sequenza ci si aspetta l'inserimento di una vocale d'appoggio indipendentemente dall'ordine del luogo di articolazione.

2. *Metodo*

Le sequenze oggetto di indagine, che riguardano le fricative alveolari e post-alveolari /sʃ - ʃs - sʒ - ʒs - zʃ - zʒ/, sono state studiate all'interno del contesto vocalico /a_i/, al confine di parole inserite all'interno di una frase cornice. Gli stimoli sono stati osservati variando la struttura prosodica, poiché, come già specificato, la presenza di un confine può interferire con il processo coarticolatorio riducendone la sovrapposizione (Byrd, Choi, 2006), e la velocità di eloquio, poiché una velocità sostenuta facilita la coarticolazione (Byrd, Tan, 1996). Per quanto riguarda la struttura prosodica, sono stati creati due contesti, inserendo gli stimoli all'interno di una frase con: 1) confine di sintagma fonologico (es. *Il dit t**asse chi**noise rapidement* – Dice tazza cinese rapidamente) e con 2) confine intonativo (es. *D'abord il a dit t**asse. Chi**noise l'a dit après*" – Prima ha detto tazza. Cinese l'ha detto dopo). A velocità sostenuta di eloquio e in caso di confine fonologico è possibile attendersi una ristrutturazione, mentre nessuna ristrutturazione è prevista quando sia presente un confine intonativo (cfr. fonologia prosodica, Nespor, Vogel, 1986).

In una fase preliminare, la naturalezza di tutte le frasi è stata verificata da 4 parlanti nativi (due studenti Erasmus provenienti da Nantes di età di 24 e 25 anni e due lettori di francese della facoltà di lingue dell'Università del Salento). Successivamente, tre studentesse (PI1, PI2 e PI3) della facoltà di Lingue e Letterature straniere dell'Università del Salento apprendenti di francese avanzato L2[1] (PI1, PI2, PI3) e una studentessa Erasmus di Nantes (PF4)[2] hanno letto per 7 volte il corpus, una volta a velocità normale e una volta a velocità sostenuta (per un totale di 168 frasi a parlante= 6 sequenze x due

[1] Al momento della raccolta dati, PI2 svolgeva un dottorato in linguistica francese presso l'Università del Salento, mentre PI3 aveva perseguito l'abilitazione SSIS in francese e collaborava presso la stessa Università. Il soggetto PI1 frequentava il primo anno di specialistica presso l'Università del Salento. Tutti e tre i soggetti avevano effettuato un soggiorno all'estero di sei mesi per Erasmus. Il parlante PI1 è stato a Nantes, il parlante PI2 è stato a Reims ritornando più volte in Francia per ricerche e PI3 si è recata a Saarbrücken (in Germania, ma vicino al confine francese), dove ha praticato comunque la lingua francese.

[2] Nelle varietà di francese dell'area settentrionale lo schwa inserito non ha una rappresentazione sottostante per cui o è assente o mostra una grande variabilità (Andreassen, Durand & Racine, 2016; Lyche, 2016), al contrario di quanto si riscontra per le varietà di francese dell'area meridionale, in cui lo schwa è mantenuto (Andreassen et al., 2016). Il nostro parlante nativo PF4 proviene dall'area settentrionale per cui lo schwa non ha una rappresentazione specificata. Ai fini del nostro studio questo aspetto

contesti prosodici x 2 velocità di eloquio x 7 ripetizioni). I materiali acustici e articolatori sono stati raccolti presso il CRIL (Università del Salento). I dati articolatori sono stati registrati utilizzando l'articulografo AG-500, posizionando i sensori sulle parti di interesse: asse sagittale della lingua (4); labbro inferiore e superiore (2); incisivi superiori e inferiori (2) e dietro le orecchie (2). Il segnale audio è stato registrato simultaneamente con una scheda audio Edirol a 44.1kHz. I dati sono stati analizzati dal punto di vista percettivo, acustico e articolatorio utilizzando scripts di PRAAT (Boersma, Weenink, 2010) e MatLab (Sigona, Stella, Grimaldi & Gili Fivela, 2015).

2.1 Analisi percettiva

Un'analisi percettiva è stata effettuata dai due autori per verificare, in modo autonomo, se il confine prosodico atteso fosse stato realizzato e la realizzazione delle sequenze e fenomeni (v. presenza di vocale epentetica).

2.2 Analisi acustica e misurazioni

L'etichettatura acustica ha riguardato i segmenti della sequenza $V_1C_1\#C_2V_2$, (ad es. /as/#/ʃi/) includendo una possibile vocale d'appoggio (V0) e/o una pausa (P) al confine tra le due consonanti. Le misurazioni hanno riguardato: i) durata di ciascun segmento; ii) durata della sequenza; iii) F1 e F2 per /a/, /i/ e l'eventuale vocale d'appoggio V0, calcolate nella porzione stabile della vocale, cioè nella porzione in cui i correlati acustici non variavano (si assumeva dunque che la conformazione linguale avesse raggiunto il *target* e l'influenza del contesto adiacente fosse minima). Inoltre sono state calcolate: iv) velocità di eloquio e di articolazione, ossia il rapporto tra la durata della sequenza e il numero di sillabe prodotte (includendo le pause per calcolare la velocità di eloquio ed escludendole per ottenere la velocità di articolazione).

2.3 Analisi articolatoria e misurazioni

L'etichettatura articolatoria ha riguardato il gesto di chiusura e apertura, sia per quanto riguarda i dati di posizione che per i picchi di velocità. Sono state considerate le traiettorie relative ai:

- movimenti della punta della lingua (*Tongue Tip*, TT), asse verticale (asse z) e orizzontale (asse x), per entrambe le fricative (alveolare e postalveolare), essendo consonanti coronali;
- movimenti del dorso della lingua (*Tongue Dorsum*, TD), asse verticale (asse z) e orizzontale (asse x), per il passaggio vocalico V-to-V ([a]- [i]);
- movimenti del labbro inferiore (*Lower Lip*, LL), lungo orizzontale (asse x), per la protrusione della fricative postalveolare, sebbene anche le fricative alveolari possano presentare un proprio gesto di protrusione (Perkell, 1986 in Kühnert, Nolan, 1999; Engwall, 2000).

---

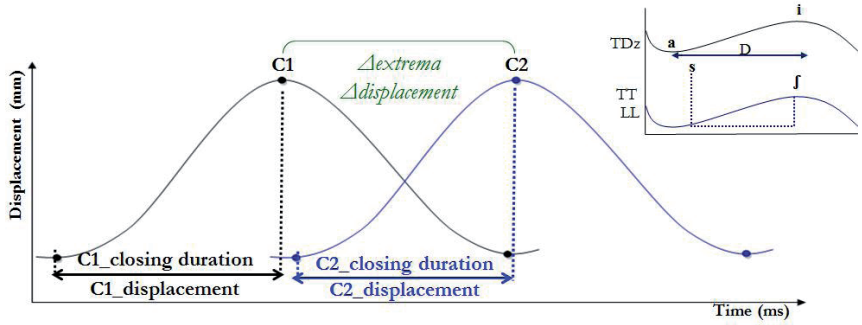
risulta particolarmente utile, poiché ci permette di confrontare e capire meglio le realizzazione degli apprendenti anche sulla base della variabilità fonetica nelle produzione del parlante nativo.

Durante la fase di etichettatura si è osservato principalmente il dorso della lingua per individuare i *target* relativi alle vocali, essendo l'articolatore maggiormente coinvolto per la loro realizzazione. Per le fricative sono state osservate, invece, le traiettorie della punta della lingua e del labbro inferiore.

Sono state calcolate successivamente le seguenti misurazioni:

– durata (ms) e ampiezza (mm) del gesto di chiusura per ciascuna fricativa (Figura 1: *C1/C2 closing_duration; C1/C2 closing_displacement*) (Byrd, Kaun, Narayanan & Saltzman, 2000);
– durata intervallo (ms) e ampiezza (mm) tra il *target* della seconda consonante (C2) e il target della prima consonante (C1) (Figura 1: *Δextrema; Δdisplacement*) (Byrd et al., 2000);
– fase relativa (ms): il rapporto tra l'intervallo tra C2-C1 e la durata del passaggio vocalico [a]-[i] (Tiede, Shattuck-Hufnagel, Johnson, Ghosh, Mattheis, Zandipur & Perkell, 2007) (Figura 1: riquadro in alto a destra).

Figura 1 - *Schema delle misurazioni articolatorie*



Le analisi statistiche sono state effettuate in SPSS. Le variabili in esame non presentano una distribuzione normale per cui si è scelto di procedere con i test non parametrici Kruskal-Wallis e Mann-Whitney, effettuando le analisi separatamente per soggetto. In questo lavoro, saranno presentate solo le analisi statistiche relative alla differenza di durata tra i *target* delle due fricative (*Δextrema*) al fine di osservare come i due gesti si coordinano in base alla presenza o meno di una vocale d'appoggio. Questa misurazione è rappresentata nella figura in basso. Tale misurazione è stata calcolata relativamente allo stesso articolatore e asse quando erano presenti entrambi i *target* consonantici (ad esempio, entrambi realizzati grazie alla punta della lingua TT o al labbro inferiore LL), in altri casi si è proceduto incrociando le informazioni o all'interno dello stesso articolatore guardando assi diversi (ad esempio punta della lingua asse verticale TTz e punta della lingua asse orizzontale TTx) o tra due articolatori diversi (ad esempio punta della lingua e labbro inferiore).

Figura 2 - *Rappresentazione grafica delle misurazioni articolatorie*

Prima di passare ai risultati dei dati articolatori, presentiamo i risultati relativi alla frequenza con la quale i parlanti inseriscono una vocale d'appoggio, e ai risultati sulla durata e qualità di tale vocale.

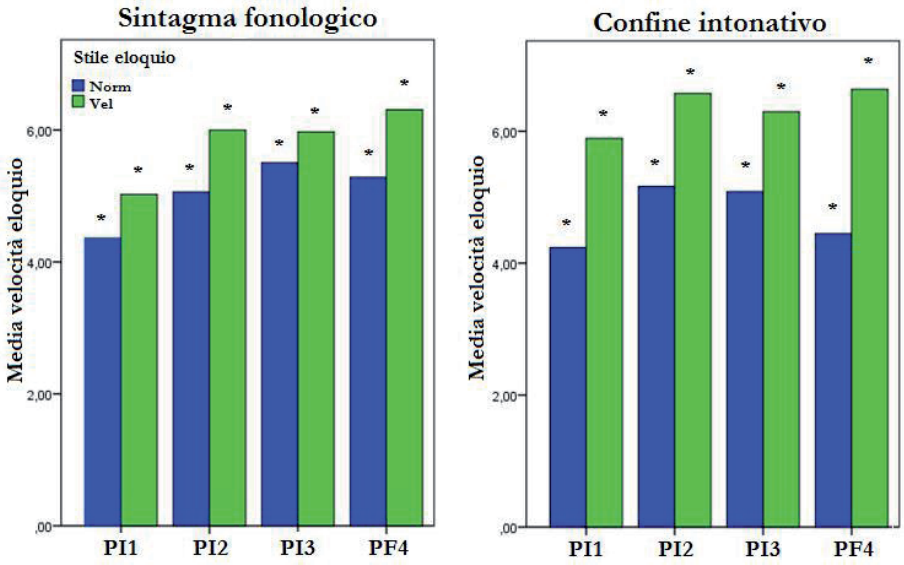
3. *Risultati*

3.1 Risultati acustici

3.1.1 Velocità di eloquio

Tutti i parlanti hanno effettivamente variato lo stile di eloquio Il test di Wilcoxon, infatti, indica una differenza statistica per tutti i parlanti, sia per le produzioni con sintagma fonologico [PI1: Z=–5.256 p=,000; PI2: Z=–5.287 p=,000; PI3: Z=–3.589 p=,000; PF4: Z=–5.309 p=,000; PF5: Z=–5.483 p=,000] che con confine intonativo [PI1: Z=–5.232 p=,000; PI2: Z=–5.086 p=,000; PI3: Z=–5.579 p=,000; PF4: Z=–5,656 p=,000; PF5: Z=–5.590 p=,000].

Figura 3 - *Velocità di eloquio in caso di sintagma fonologico (a sinistra) e di confine intonativo (a destra) per tutti i parlanti*



La velocità di articolazione è stata calcolata per confrontare le produzioni degli apprendenti con quella del nativo. Una serie test di Mann-Whitney, effettuando

confronti incrociati (velocità normale degli apprendenti *vs* velocità normale del francofono; velocità sostenuta degli apprendenti *vs* velocità normale e sostenuta del francofono), mostra una differenza significativa (p=,000).

3.1.2 Frequenza dell'inserimento di vocale

A velocità normale di eloquio, tutti i parlanti inseriscono una vocale tra le due consonanti, per entrambe le condizioni prosodiche e per le due sequenze. Le tabelle riportate di seguito mostrano le frequenze relative all'inserimento di una vocale per gli apprendenti (Tabella 1) e per il parlante nativo (Tabella 2). Come si può osservare, tra gli apprendenti solo il parlante PI3 realizza alcuni casi in cui non inserisce alcuna vocale, per la sequenza AP in assenza di confine prosodico (5 casi su 30). PF4, invece, in presenza di confine intonativo realizza sempre una pausa che può essere preceduta o meno da una vocale.

Tabella 1 – *Frequenze relative all'inserimento di una vocale d'appoggio da parte degli apprendenti italofoni a velocità normale di eloquio e nelle due condizioni prosodiche*

Velocità normale di eloquio – Apprendenti italofoni							
Sintagma fonologico							
Alveolare-postalveolare				*Postalveolare-alveolare*			
P	Schwa	Tot	%	P	Schwa	Tot	%
PI1	28	28	100	PI1	14	14	100
PI2	29	29	100	PI2	14	14	100
PI3	25	30	83,3	PI3	14	14	100
Confine intonativo							
PI1	28	28	100	PI1	13	13	100
PI2	27	27	100	PI2	14	14	100
PI3	26	26	100	PI3	14	14	100

Tabella 2 - *Frequenze relative all'inserimento di una vocale d'appoggio da parte del parlante nativo a velocità normale di eloquio e nelle due condizioni prosodiche*

Velocità normale di eloquio – parlante nativo							
Sintagma fonologico							
Alveolare-postalveolare				*Postalveolare-alveolare*			
P	Schwa	Tot	%	P	Schwa	Tot	%
PF4	28	28	100	PF4	14	14	100
Confine intonativo							
PF4	16	28	57,1	PF4	5	14	35,7

A velocità sostenuta, tra gli apprendenti si distingue PI1 che, come mostra la Tabella 3, per la sequenza AP in presenza di confine intonativo realizza una vocale all'interno del nesso solo in 3 casi. In tutti gli altri contesti, PI1 realizza assimilazioni di luogo. I parlanti PI2 e PI3, al contrario, continuano a inserire una vocale soprattutto in presenza di confine intonativo. Come mostra la Tabella 4, invece, PF4 produce una vocale solo per la sequenza AP in caso di un confine intonativo, in tutti gli altri casi realizza assimilazioni di luogo[3].

Tabella 3 - *Frequenze relative all'inserimento di una vocale d'appoggio da parte degli apprendenti italofoni a velocità sostenuta di eloquio e nelle due condizioni prosodiche*

Velocità sostenuta di eloquio – Apprendenti italofoni							
Sintagma fonologico							
Alveolare-postalveolare				*Postalveolare-alveolare*			
Spk	Schwa	Tot	%	Spk	Schwa	Tot	%
PI1	0	26	–	PI1	0	12	–
PI2	7	30	23,3	PI2	8	17	47,1
PI3	9	29	31,0	PI3	8	14	57,1
Confine intonativo							
PI1	3	24	12,5	PI1	0	11	–
PI2	12	24	50	PI2	7	11	63,6
PI3	16	27	59,3	PI3	11	14	78,6

Tabella 4 - *Frequenze relative all'inserimento di una vocale d'appoggio da parte del parlante nativo a velocità sostenuta di eloquio e nelle due condizioni prosodiche*

Velocità sostenuta di eloquio – parlante nativo							
Sintagma fonologico							
Alveolare-postalveolare				*Postalveolare-alveolare*			
Spk	Schwa	Tot	%	Spk	Schwa	Tot	%
PF4	0	28	-	PF4	0	13	-
Confine intonativo							
PF4	12	20	42,9	PF4	0	13	-

[3] Per i risultati sulla realizzazione delle assimilazioni di luogo da parte di questi soggetti si rimanda ai lavori di d'Apolito (2012) e d'Apolito, Gili Fivela (2013).

3.1.3 Durata del segmento

A velocità normale di eloquio e in assenza di confine prosodico, la durata del segmento vocalico si colloca con maggiore frequenza tra i 50 ms e i 110 ms, mentre in presenza di confine intonativo si ha una maggiore variabilità per cui l'intervallo di durata oscilla tra i 70-150 ms per PI1 e PI2 e tra 50-110 ms per PI3 e PF4. A velocità sostenuta di eloquio, per PI2 e PI3 la vocale ha una durata tra i 30 ms e gli 80 ms in entrambe le condizioni prosodiche. Per PI1, nei pochi casi in cui inserisce una vocale, si osserva una durata variabile tra 20-100 ms; per il parlante francofono (PF4), il range è, invece, tra 50-140 ms.

Figura 4 - *Durata del segmento vocalico (ms) a velocità normale (a sinistra) e sostenuta (a destra) di eloquio, nelle due condizioni prosodiche per tutti i parlanti*

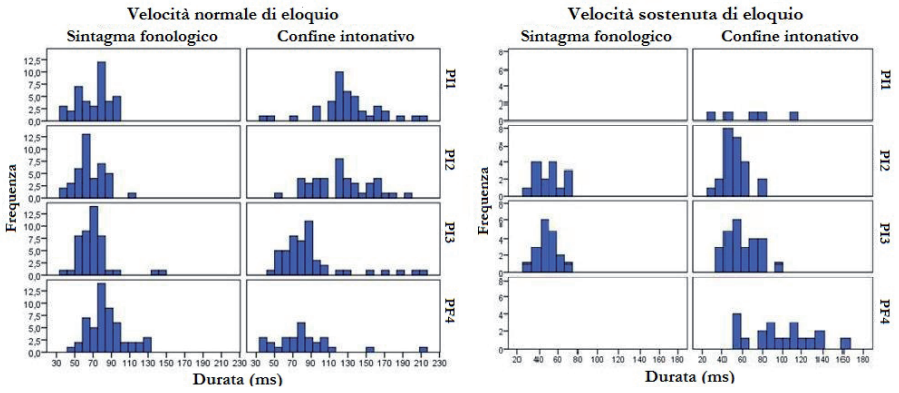
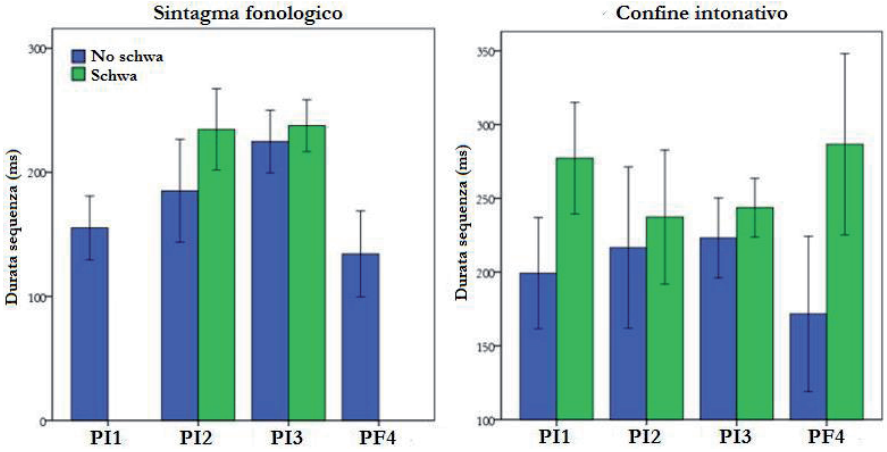


Tabella 5 - *Valori medi (in ms) e deviazioni standard relativi alla durata del segmento vocalico V0 in caso di sintagma fonologico (a sinistra) e di confine intonativo (a destra) per tutti i parlanti a velocità normale (sopra) e sostenuta (sotto) di eloquio*

Velocità normale			
Sintagma fonologico		*Confine intonativo*	
Spk	Media (dev. st.)	Spk	Media (dev. st.)
PI1	66.83 (22.39)	PI1	130.46 (35.11)
PI2	68.12 (21.43)	PI2	121.75 (32.24)
PI3	64.45 (25.24)	PI3	91.35 (42.29)
PF4	82.25 (21.90)	PF4	42.52 (48.47)
Velocità sostenuta			
PI1	0	PI1	5.48 (17.64)
PI2	20.61 (26.83)	PI2	28.62 (26.85)
PI3	20.20 (24.75)	PI3	37.54 (31.22)
PF4	0	PF4	49.74 (53.34)

Per le produzioni a velocità di eloquio sostenuta è stato anche osservato se la presenza o meno di una vocale d'appoggio contribuisse alla maggiore durata della sequenza. Come si può osservare dal grafico in Figura 5, nel caso in cui venga realizzata una vocale tra le due fricative, la durata della sequenza è maggiore rispetto alla sequenza senza vocale. Un T-test a campioni indipendenti indica che in assenza di confine prosodico la significatività è raggiunta solo per PI2 ($t_{(39)}$=-3,680; p=,001). In presenza di confine intonativo, la differenza è significativa per PI1 ($t_{(40)}$=-4,688; p=,000), per PI3 ($t_{(40)}$=-2,825; p=,000) e per PF4 ($t_{(40)}$=-6,513; p=,000).

Figura 5 - *Durata della sequenza (ms) a velocità sostenuta di eloquio in base alla presenza (verde) e assenza (blu) del segmento vocalico*



3.1.4 Formanti F1 e F2

Al fine di osservare le caratteristiche qualitative del segmento vocalico realizzato tra le due fricative, ne sono state calcolate le prime due formanti. I grafici scatterplot riportati in basso rappresentano le aree di esistenza delle vocali realizzate all'interno della sequenza *target*, cioè [a] (in blu), [i] (in verde) e V0 (in rosso), a velocità di eloquio normale (Figura 6) e sostenuta (Figura 7) e per entrambe le condizioni prosodiche. Come si può osservare, per gli apprendenti italofoni l'area di esistenza della vocale d'appoggio è molto variabile e, almeno per alcuni parlanti, sembra assuma caratteristiche diverse a seconda del contesto prosodico nel quale viene realizzata: per il francofono è qualitativamente più circoscritta. Inoltre, la vocale realizzata dagli apprendenti ha le caratteristiche di una vocale medio-alta o alta, più anteriore di [a] (bassa F1 e alta F2), mentre per il francofono mostra caratteristiche maggiormente simili a una vocale centralizzata.

Figura 6 - *F1 e F2 delle vocali realizzate a velocità normale e per entrambe le condizioni prosodiche (sint = sintagma fonologico; inton = confine intonativo):/a/ in blu; /i/ in verde e V0 in rosso*

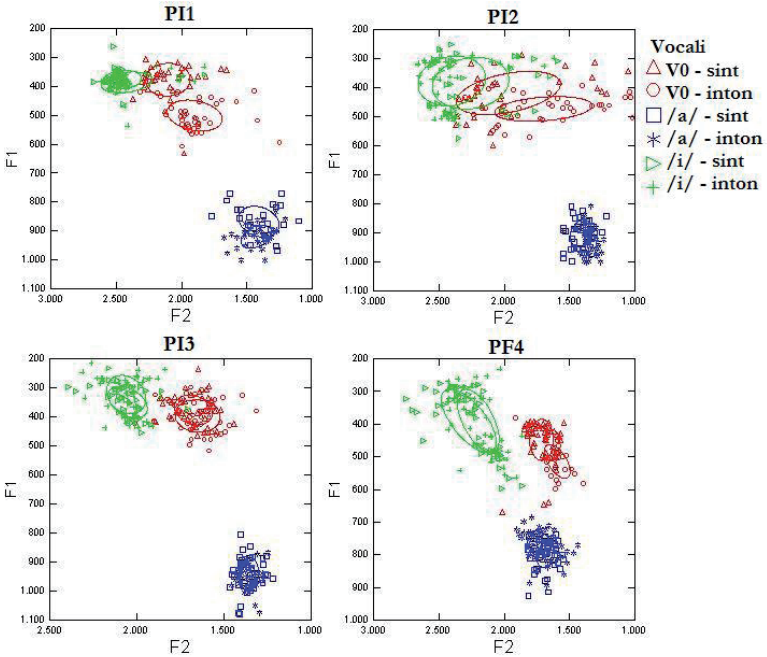
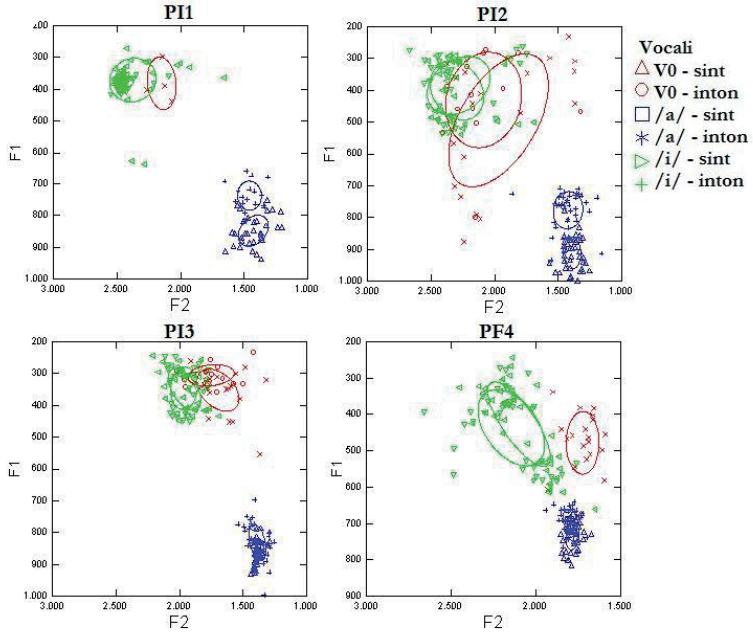


Figura 7 - *F1 e F2 delle vocali realizzate a velocità sostenuta e per entrambe le condizioni prosodiche (sint = sintagma fonologico; inton = confine intonativo):/a/ in blu; /i/ in verde e V0 in rosso*



Quindi, riassumendo, dal punto di vista acustico, a velocità normale di eloquio, tutti i parlanti realizzano un segmento vocalico tra le due consonanti, mentre a velocità sostenuta il numero di vocali diminuisce notevolmente e, di fatto, solo due parlanti italofoni su tre (PI2 e PI3) continuano a inserire una vocale all'interno del nesso, con una frequenza maggiore in caso di confine intonativo. L'altro parlante italofono (PI1) riduce invece drasticamente il numero di casi di inserzione vocalica, realizzando perlopiù assimilazioni di luogo e comportandosi, da questo punto di vista, in modo simile al parlante francofono. La durata del segmento è variabile per tutti parlanti e si attesta tra 50-150 ms e, a velocità sostenuta di eloquio, l'inserzione della vocale contribuisce ad una maggiore durata della sequenza. Per gli apprendenti anche la qualità della vocale, in termini di F1 e F2, è molto variabile e si presenta come una vocale più chiusa e anteriore rispetto a un canonico schwa. Per il parlante nativo, invece, la vocale mostra meno variabilità e presenta le caratteristiche di una tipica vocale schwa. Queste caratteristiche fanno pensare, rispetto a quanto descritto in letteratura, che la vocale inserita dagli apprendenti italofoni sia una vocale intrusiva, benché si possa comunque considerare che svolga la funzione di una vocale epentetica, ossia favorisca la produzione di un nesso marcato.

3.2 Risultati articolatori

Un risultato importante riguarda il fatto che sul dorso della lingua sono stati sempre individuati, per tutti i parlanti, i *target* relativi alle due vocali estreme della sequenza, cioè [a] e [i], mentre non è stato individuato alcun *target* per la vocale intermedia. Anche nel caso sia udibile una vocale, quindi, tra i bersagli articolatori dei due gesti consonantici si realizza solo un gesto di apertura.

A velocità normale di eloquio, il gesto di apertura si realizza in modo sistematico per entrambe le sequenze e in entrambe le condizioni prosodiche. L'apprendente PI3 realizza alcuni casi (5) senza un vero gesto di apertura tra le due fricative e senza alcun *output* acustico per la sequenza AP (/zʃ/e /zʒ/) in assenza di confine prosodico. Considerando il *Δextrema*, cioè la differenza tra i due target consonantici, il test di Kruskal-Wallis indica una differenza significativa su TTz [$\chi 2(1,29)=5,602$; p=,018] e LL-TTz [$\chi 2(1,29)=6,770$; p=,009]. I valori medi indicano una durata maggiore per i casi in cui sia stata inserita una vocale (165ms) rispetto ai casi senza *output* acustico e gesto di apertura (125ms). In caso di un confine intonativo, il francofono si distingue dagli italofoni, poiché inserisce sempre una pausa che può essere preceduta o meno da schwa. Tuttavia, il test di Kruskal-Wallis non indica alcuna significatività nel confronto delle misure articolatorie effettuate circa le sue produzioni, per cui l'intervallo tra i bersagli delle due fricative non cambia se interviene un solo elemento (pausa) o due elementi (pausa + schwa).

A velocità sostenuta di eloquio si ha una maggiore variabilità. In assenza di confine prosodico, PI1 e PF4 non inseriscono alcuna vocale e realizzano sequenze senza un gesto di apertura intermedio. Gli apprendenti PI2 e PI3, invece, realizzano i nessi con due o tre *pattern* articolatori: 1) gesto di apertura tra C_1 e C_2 con inserimento

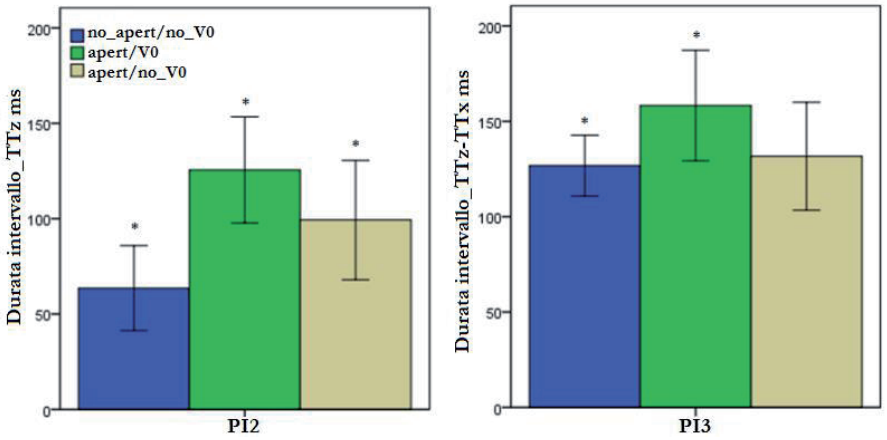
di vocale udibile (apertura/V0); 2) gesto di apertura senza *output* acustico (apertura/no V0); 3) nessun gesto di apertura e nessun *output* acustico (no apertura/no V0). Nella Tabella 6 sono riportate le frequenze di realizzazione dei tre *pattern* articolatori per entrambe le sequenze. Come si può osservare, la sequenza AP viene realizzata con tre *pattern*, mentre la sequenza PA viene realizzata con due *pattern* (casi descritti in 1 e 2).

Tabella 6 - *Frequenze relative alla realizzazione dei tre pattern articolatori a velocità sostenuta e in assenza di confine prosodico da parte di tutti i parlanti*

Sintagma fonologico									
Alveolare-postalveolare					*Postalveolare-alveolare*				
P	Casi	N	Tot	%	P	Casi	N	Tot	%
PI1	No apert/no V0	26	26	100	PI1	No apert/no V0	12	12	100
PI2	Apert/V0	7	30	23,3	PI2	Apert/V0	8	17	47,1
	Apert/no V0	14		46,7		Apert/no V0	9		52,9
	No apert/no V0	9		30		No apert/no V0	0		0
PI3	Apert/V0	9	29	31,0	PI3	Apert/V0	8	14	57,1
	Apert/no V0	5		17,2		Apert/no V0	6		42,9
	No apert/no V0	15		51,8		No apert/no V0	0		0
PF4	No apert/no V0	28	28	100	PF4	No apert/no V0	13	13	100

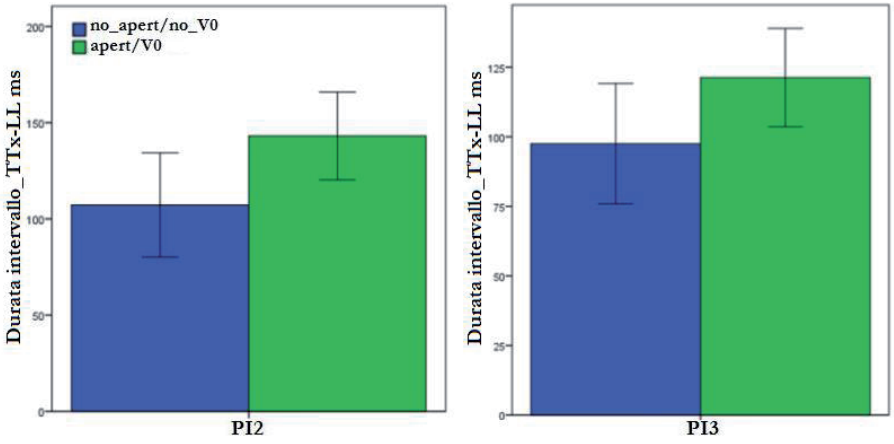
Per la sequenza AP, il test indica una differenza significativa su TTz per PI2 [χ2(2,30)=12,932; p=,002] e su TTz-TTx per PI3 [χ2(2,29)=7,609; p=,022]. Il test di Mann-Whitney mostra che per PI2 i casi no_apertura/no_V0 si distinguono dai casi apertura/V0 [Z=-3,228; p=,000] e dai casi apertura/no_V0 [Z=-2,521; p=,011]. L'intervallo medio per i casi no_apertura/no_V0 è di 63,54 ms (dev. st. 22,26), mentre per i casi apertura/V0 è di 125,54 ms (dev. st. 27,38) e per i casi apertura/no_V0 è di 99,28 ms (dev. st. 31,31). Per i PI3 il test di Mann-Whitney [Z=–2,772; p=,004] mostra una differenza significativa solo tra i casi apertura/V0 (158,26 ms, dev. st. 28,99) e i casi no_apertura/no_V0, (126,75 ms, dev. st. 15,98). Il *pattern* apertura/V0 mostra l'intervallo maggiore, a cui segue il *pattern* apertura/no_V0 e infine il *pattern* no_apertura/no_V0 con l'intervallo più breve.

Figura 8 - *Grafico a barre per il Δextrema per la sequenza alveolare-postalveolare in assenza di confine prosodico per PI2 su TTz (a sinistra) e PI3 su TTz-TTx (a destra)*



Per la sequenza PA, il test di Kruskal-Wallis è significativo su TTx-LL per PI2 [$\chi2(1,17)=5,346$; p=,021] e PI3 [$\chi2(1,14)=4,067$; p=,044]. In particolare, l'intervallo è maggiore per i casi apertura/V0 (PI2: 143,12 ms, dev. st. 22.,82; PI3: 121,25 ms, dev. st. 17,67) rispetto ai casi apertura/no_V0 (PI2: 107,25 ms, dev. st. 27,09; PI3: 97,50 ms, dev. st. 21,62).

Figura 9 - *Grafico a barre per il Δextrema per la sequenza postalveolare-alveolare in assenza di confine prosodico per PI2 (a destra) e PI3 s (a sinistra) su TTx-LL*



Per quanto riguarda i risultati relativi alle sequenze in presenza di confine intonativo, come si può osservare dalla Tabella 7 si realizzano tre *pattern* per la sequenza AP e due *pattern* per la sequenza PA, come nel contesto precedente. Il madrelingua PF4 realizza la sequenza inserendo uno schwa o assimilando per il luogo di articolazione. Il parlante PI1, comportandosi in modo simile, tende a non inserire una vocale tra

le due fricative realizzando, quindi, la sequenza con o senza gesto di apertura. PI2 e PI3, invece, realizzano sempre il gesto di apertura accompagnato o meno da *output* acustico.

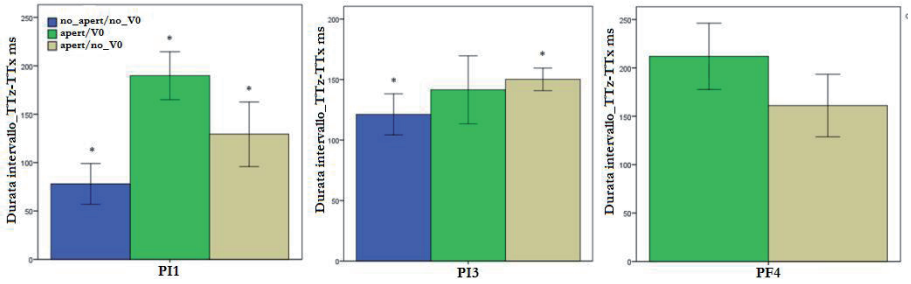
Tabella 7 - *Frequenze relative alla realizzazione dei tre pattern articolatori a velocità sostenuta e in presenza di confine intonativo da parte di tutti i parlanti*

Confine intonativo									
Alveolare-postalveolare					*Postalveolare-alveolare*				
P	Casi	N	Tot	%	P	Casi	N	Tot	%
PI1	Apert/V0	4	24	16,66	PI1	Apert/V0	0	11	0
	Apert/no V0	12		50		Apert/no V0	5		45,5
	No apert/no V0	8		33.34		No apert/no V0	6		54,5
PI2	Apert/V0	13	24	54,16	PI2	Apert/V0	7	11	63,6
	Apert/no V0	8		33,34		Apert/no V0	4		36,4
	No apert/no V0	3		12,5		No apert/no V0	0		0
PI3	Apert/V0	15	27	55,56	PI3	Apert/V0	11	14	78,6
	Apert/no V0	4		14,82		Apert/no V0	3		21,4
	No apert/no V0	8		29,62		No apert/no V0	0		0
PF4	Apert/V0	13	28	46.42	PF4	Apert/V0	7	14	50
	Apert/no V0	7		25		Apert/no V0	0		0
	No apert/no V0	8		28,58		No apert/no V0	7		50

Per la sequenza AP, il test di Kruskal-Wallis sui valori del *Δextrema* raggiunge la significatività su TTz per PI1 [$\chi2(2,24)=18,184$; p=,000] e PI3 [$\chi2(2,27)=10,268$; p=,006]. Il test di Mann-Whitney indica che per PI1 i casi no_apertura/no_V0 si distinguono dai casi apertura/V0 [Z=−2,717; p=,004] e dai casi apertura/no_V0 [Z=−3,626, p=,000] poiché i casi senza gesto di apertura e senza *output* acustico corrispondono all'intervallo più breve (71,69 ms; dev. st.15,39) rispetto ai casi apertura/no_V0 (120,55 ms; dev. st. 28,84) e apertura/V0 (186,05 ms; dev. st. 22,56). Inoltre, si riscontra anche una differenza significativa tra i casi apertura/V0 e i casi apertura/no_V0 [Z=−2,668, p=,008] con valori medi di 186,05 ms (dev. st. 22,56) e 120,55 ms (28,84) rispettivamente. Per il parlante PI3, il test indica una differenza tra no_apertura/no_V0 e apertura/V0 (Z=−3,099; p=,001). I casi con inserimento di vocale corrispondono invece ad un intervallo maggiore (150,55 ms; dev. st. 23,92) rispetto ai casi no_apertura/no_V0 (107,46 ms, dev. st. 25,44). I casi con apertura intermedia e senza *output* acustico hanno una durata di circa 145ms (dev. st. 13,91), quindi simile a quella dei casi in cui è stato effettivamente inserita una vocale. Per il parlante francofono, l'analisi ha riguardato solo i casi in cui fosse presente il gesto di apertura, con o senza output acustico: i casi con *output* acustico, hanno un intervallo maggiore (200,64 ms; dev. st. 30,70) rispetto ai casi con ge-

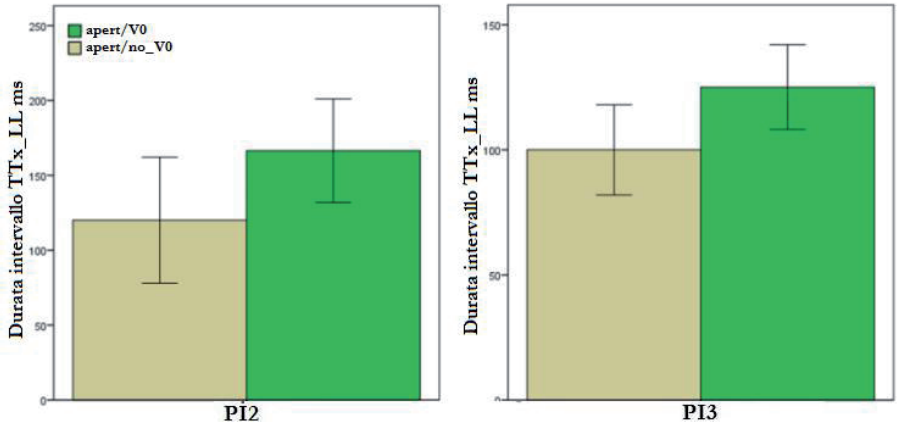
sto di apertura e senza che si percepisca una vocale (142,74 ms; dev. st. 36,77) PF4 [$\chi 2(1,20)=8,391$; p=,004.

Figura 10 - *Grafico a barre per il Δextrema per la sequenza alveolare-postalveolare in presenza di confine intonativo su TTz-TTx per PI2 (a sinistra), PI3 (centro), PF4 (a destra)*



Infine, come abbiamo già detto, gli apprendenti PI2 e PI3 realizzano la sequenza PA sempre con un gesto di apertura che può avere o meno un *output* acustico. Per PI2 questo corrisponde ad una differenza significativa in termini di *Δextrema* su TTz [$\chi 2(1,11)=6,036$; p=,014], mentre per PI3 su TTx [$\chi 2(1,14)=6,303$; p=,012]. Per entrambi i parlanti, l'intervallo è maggiore nel caso in cui ci sia *output* acustico rispetto ai casi in cui sia stato realizzato solo un gesto di apertura (rispettivamente per PI2: 149,28 ms dev. st. 7,86 vs 124,08 ms; dev. st. 14,99; per PI3 167,77 ms dev. st. 17,60 vs. 133,33 dev. st. 5,77). Anche PI1 realizza la sequenza con o senza gesto di apertura, ma questo non corrisponde ad una differenza significativa delle misure da noi considerate. Il francofono PF4 inserisce una vocale o assimila per il luogo di articolazione. Trattandosi di *pattern* articolatori corrispondenti a strategie che riguardano segmenti diversi non si è ritenuto opportuno procedere con un'analisi statistica.

Figura 11 - *Grafico a barre per il Δextrema per la sequenza postalveolare-alveolare in presenza di confine intonativo su TTx-LL per PI2 (a sinistra) e PI3 (a destra)*



Riassumendo, quindi, a velocità normale di eloquio, il *pattern* articolatorio realizzato da tutti i parlanti corrisponde a un gesto di apertura tra le due fricative, senza un *target* specifico per la vocale inserita benché il segmento vocalico risulti sul piano acustico e percettivo. A velocità sostenuta si ha una maggiore variabilità. Si individuano tre tipi di realizzazione: 1) gesto di apertura tra C_1 e C_2 con inserimento di vocale (apertura/V0); 2) gesto di apertura senza *output* acustico (apertura/no V0); 3) nessun gesto di apertura e nessun *output* acustico (no apertura/no V0). La frequenza delle tre realizzazioni varia in base al parlante e al tipo di sequenza. Generalmente, si riscontra una differenza significativa nella misura del *Δextrema* tra i casi in cui si ha il gesto di apertura con *output* acustico, con l'intervallo più lungo, e i casi senza gesto di aperura e senza *output* acustico, con intervallo più breve. Quando è stato realizzato un gesto di apertura senza segmento vocalico, invece, la durata è intermedia e differisce dai casi con gesto di apertura e *output* vocalico; si hanno, invece, poche evidenze rispetto ai casi senza gesto di apertura e senza *output* acustico.

4. *Discussione e conclusioni*

In questo studio si è osservata la realizzazione di sequenze di sibilanti, alveolare-postalveolare e postalveolare-alveolare, da parte di apprendenti italofoni di francese L2 avanzato, focalizzando l'attenzione, dal punto di vista acustico e articolatorio, sulla vocale d'appoggio che viene realizzata al fine di riparare nessi consonantici non nativi. Nell'articolo, è stato sottolineato che l'inserzione vocalica può realizzarsi o come epentesi vocalica, cioè una vocale che presenta un proprio *target* articolatorio, oppure come vocale intrusiva, senza alcun *target* articolatorio dovuta ad un gesto di transizione tra le due consonanti.

In questa sede, l'inserzione della vocale è stata studiata prendendo in considerazione due fattori che possono interferire con la realizzazione dei nessi consonantici e possono quindi avere un effetto sulla realizzazione della vocale d'appoggio: l'influenza della velocità di eloquio e della presenza di confini prosodici diversi. In generale, l'inserzione vocalica si riscontra in modo sistematico a velocità normale di eloquio, in entrambe le condizioni prosodiche, per entrambe le sequenze e per tutti i parlanti. Le sole eccezioni si hanno per un apprendente, PI3, che realizza pochi casi senza una vocale di appoggio e per PF4 che, in presenza di confine intonativo, realizza sempre una pausa preceduta o meno dal segmento vocalico. A velocità sostenuta di eloquio, il segmento vocalico viene inserito soprattutto da due apprendenti (PI2 e PI3). L'apprendente PI1 tende a non inserire una vocale a velocità sostenuta, di fatto avvicinandosi al comportamento del francofono, che realizza una vocale solo in caso di confine intonativo e per la sequenza alveolare-postalveolare.

Dal punto di vista acustico, sono state osservate la durata e i valori della prima e seconda formante. La durata del segmento vocalico è variabile per tutti i parlanti e incide sulla durata maggiore dell'intera sequenza. Per gli italofoni la vocale inserita

non mostra uno spazio fonetico ben definito, né centralizzato, tale da far pensare a un canonico schwa, poiché il segmento inserito risulta essere una vocale abbastanza chiusa e più anteriore di [a] (sembra influenzata dalla vocale successiva [i]). Per il francofono, invece, il segmento vocalico mostra uno spazio fonetico definito con le caratteristiche centralizzate di una vocale schwa. Il segmento inserito, quindi, nel presentare durata e qualità variabile pare influenzato dallo stile di eloquio per cui, secondo quanto descritto in letteratura, tale vocale potrebbe essere considerata una vocale intrusiva. Tuttavia, tenuto conto del fatto che per gli apprendenti italofoni l'inserzione vocalica è supposta avere la funzione di riparare il nesso marcato nella loro L1, è verosimile pensare comunque che la funzione svolta sia quella di una vocale epentetica. D'altronde, bisogna considerare che l'elemento vocalico si inserisce in posizione di coda in una sillaba non accentata, posizione che già di per sé può favorire un'alta variabilità in durata e qualità poiché fortemente influenzata dal contesto adiacente. Poiché dal punto di vista acustico è parso piuttosto difficile poter distinguere chiaramente tra vocale intrusiva e epentetica, si sono osservate anche le produzioni delle sequenze dal punto di vista articolatorio.

Osservando la traiettoria del dorso della lingua, che dà indicazioni sui gesti vocalici, non si è riscontrato alcun *target* specifico per la vocale inserita, per nessun parlante, mentre sono chiaramente visibili i *target* delle due vocali della sequenza [a] e [i]. È stato quindi calcolato l'intervallo tra i due *target* consonantici (differenza tra il *target* di C_2 e di C_1 - *Δextrema*), al fine di osservare la coordinazione tra le due fricative in base alla presenza o meno di una vocale inserita. A velocità normale di eloquio, il pattern articolatorio tipico che si è riscontrato per tutti i parlanti, per le due sequenze e per entrambe le condizioni prosodiche, è un gesto di apertura tra i due gesti consonantici. Si distingue il parlante francofono che in presenza di confine intonativo e per entrambe le sequenze realizza una pausa che può essere preceduta o meno dallo schwa. Dal punto di vista articolatorio, la coordinazione temporale tra le sue due fricative non sembra subire modificazioni poiché l'intervallo tra i due target consonantici non varia in base all'inserimento di due elementi (pausa e vocale) o di un solo elemento (vocale). Si è osservato un effetto di compensazione sulla durata della fricativa in sillaba finale (C_1) poiché si osserva una durata minore nel caso di inserimento di pausa e vocale e una durata maggiore nel caso di inserimento solo di pausa. A velocità sostenuta di eloquio, si riscontrano tre tipi di realizzazione: 1) gesto di apertura tra C_1 e C_2 con inserimento di vocale; 2) gesto di apertura senza *output* acustico; 3) nessun gesto di apertura e nessun *output* acustico. I tre tipi di realizzazione si riscontrano nelle produzioni di due apprendenti su tre (PI2 e PI3) per la sequenza alveolare-postalveolare in entrambe le condizioni prosodiche, mentre per il terzo italofono (PI1) e per il parlante francese nativo solo in presenza di confine intonativo. In ogni caso. la sequenza postalveolare-alveolare viene realizzata con due dei tre *pattern* articolatori, ossia un gesto di apertura accompagnato o meno un *output* acustico (casi descritti in 1 e 2). Tra gli apprendenti si distingue, quindi, PI1 che solo in assenza di confine prosodico non realizza mai una vocale ma assimila la sequenza per il luogo di articolazione,

come fa il francofono; in presenza di confine realizza la sequenza con o senza gesto di apertura, e comunque senza *output* acustico (casi descritti in 2 e 3). Anche il parlante francofono non realizza mai una vocale in assenza di confine prosodico, assimilando la sequenza per il luogo di articolazione, mentre in presenza di confine realizza un gesto di apertura con schwa (caso descritto in 1) o assimila. I casi con gesto di apertura e *output* acustico motivano la durata maggiore delle sequenze, distinguendosi dai casi senza gesto di apertura e segmento vocalico, con durata più breve, e dai casi con gesto di apertura senza *output* acustico, con durata intermedia.

Dai risultati emerge, quindi, chiaramente l'influenza sia della velocità di eloquio che della struttura prosodica. L'inserimento della vocale si ha maggiormente a velocità normale di eloquio, mentre a velocità sostenuta si ha una maggiore variabilità. Per le produzioni a velocità sostenuta, soprattutto per gli italofoni PI2 e PI3, si nota un maggiore numero di inserimenti, soprattutto in presenza di un confine intonativo. Per quanto riguarda il tipo di sequenza (postalveolare-alveolare e alveolare-postalveolare), la vocale d'appoggio si riscontra in entrambe le sequenza sebbene si noti che, a velocità sostenuta di eloquio e sempre per gli apprendenti PI2 e PI3, percentuali di inserimento maggiori per la sequenza postalveolare-alveolare piuttosto che per la sequenza opposta.

In sostanza, per tutti i parlanti, i risultati articolatori mostrano che il segmento vocalico inserito è di tipo "*targetless*" poiché non è stato possibile individuare un *target* specifico sulla traiettoria del dorso della lingua. I risultati relativi al segmento vocalico inserito dal parlante francofono suggeriscono che derivi proprio dal rilascio acustico della consonante (Rialland, 1986), dovuto alla velocità normale di eloquio e alla presenza di confine intonativo a velocità sostenuta. Per quanto riguarda gli apprendenti, i tre tipi di realizzazione indicano una certa variabilità articolatoria legata, verosimilmente, ai tentativi nel realizzare nessi marcati nella lingua nativa per cui l'*output* acustico sembrerebbe dovuto ad un vero e proprio *gestural mistiming*.

In conclusione, le produzioni degli apprendenti italofoni risultano qui influenzate dalle caratteristiche fonetiche e fonologiche della lingua nativa per cui, com'è noto, il fine dell'inserimento vocalico è di riparare un nesso non nativo, marcato. Questo studio suggerisce, inoltre, che l'inserimento vocalico sia dovuto a un *gestural mistiming*, quindi legato a questioni articolatorie con particolare riferimento alla scarsa coordinazione tra le due consonanti: la produzione delle consonanti creerebbe uno spazio articolatorio sufficiente perché si realizzi un segmento vocalico d'appoggio, benché questo non corrisponda ad un vero bersaglio. In futuro, oltre ad un'analisi percettiva da parte di parlanti nativi delle produzioni relativamente all'inserzione vocalica e/o alla realizzazione di altri fenomeni coarticolatori, si prevede di incrementare il campione degli apprendenti, come anche quello dei parlanti nativi, al fine di poter maggiormente confermare i risultati qui presentati.

Riferimenti bibliografici

Boersma, P., Weenink, D. (2010). PRAAT: Doing phonetics by computer. Version 5.2. htttp://www.praat.org.

Browman, C.P. (1995). Dynamics and articulatory phonology. In *Minds as motion*, 175-193.

Browman, C.P., Goldstein, L. (1986). Towards an articulatory phonology. In *Phonology Yearbook*, 3, 219-252.

Browman, C.P., Goldstein, L. (1987). Tiers in Articulatory Phonology with some implications for casual speech. In *Haskins Laboratories Status Report on Speech research*, SR-92, 1-30.

Browman, C.P., Goldstein, L. (1989). Articulatory gestures as phonological units. In *Phonology*, 6, 201-251.

Browman, C.P., Goldstein, L. (1992). "Targetless" Schwa: An Articulatory Analysis. In Docherty, G.J., Ladd, D.R. (Eds.), *Papers in Laboratory Phonology II: Gesture, segment, prosody*. Cambridge: CUP, 26-56.

Browman, C.P., Goldstein, L. (2007). The gestural phonology model, Speech production: motor control. In *Brain research and fluency disorders*, 57-71.

Byrd, D., Choi, S. (2006). At the juncture of prosody, phonology and phonetics – the interaction of phrasal and syllable structure in shaping the timing of consonant gestures. In *Proceedings of the 10th Conference on Laboratory Phonology*.

Byrd, D., Kaun, A., Narayanan, S. & Saltzman, E. (2000). Phrasal signature in articulation. In Broe, M.B., Pierrehumbert, J.B. (Eds.), *Papers in Laboratory Phonology V.* Cambridge University Press, 70-87.

Byrd, D., Tan, C.C. (1996). Saying consonant clusters quickly. In *Journal of Phonetics*, 24, 209-244.

d'Apolito, S. (2012). La coarticolazione: Studio acustico, cinematico e percettivo di sequenze di sibilanti della lingua francese nelle produzioni di studenti italofoni. Tesi di Dottorato.

d'Apolito, S., Gili Fivela, B. (2013). Place assimilation and articulatory strategies: the case of sibilant sequences in French as L1 and L2. *Interspeech*, Lyon, France.

Davidson, L. (2005). Addressing phonological questions with ultrasound. In *Clinical Linguistics & Phonetics*, 19(6/7), 619-633.

Davidson, L. (2006). Phonology, phonetics, or frequency: Influences on the production of non-native sequences. In *Journal of Phonetics*, 34, 104-137.

Davidson, L., Stone, M. (2003). Epenthesis versus gestural mistiming in consonant cluster production: An Ultrasound study. In Garding, G., Tsujimura, M. (Eds.), *Proceedings of the West Coast Conference on Formal Linguistics* (WCCFL). Somerville, MA: Cascadilla Press, 165-178.

Eckman, R.F. (2008). Typological markedness and second language phonology. In Hansen Edwards, J.G., Zampini, M.L. (Eds.), *Phonology and second language acquisition*. Philadelphia: John Benjamins, Ch. 4, 95-115.

Engwall, O. (2000). Dynamical aspects of coarticulation in Swedish fricatives: a combined EMA & EPG study. In *TMH Quarterly Status and Progress Report*, 4, KTH, Stockholm, 49-3.

Farnetani, E., Busá, M.G. (2004). Italian clusters in continuous speech. In *Proceedings of the 3rd ICSLP*, Yokohama, Japan, vol. 1, 359-362.

Gafos, A. (2002). A grammar of gestural coordination. In *Natural Language and Linguistic Theory*, 20, 269-337.

Hall, N. (2003). Gestures and segments: Vowel intrusion as overlap. PhD Dissertation, University of Massachusetts, Amherst.

Hall, N. (2006). Cross-linguistic patterns of vowel intrusion. In *Phonology*, 23, 387-429.

Hall, N. (2011). Vowel epenthesis. In van Oostendorp, M., Ewen, C.J., Hume, E. & Rice, K. (Eds.), *The Blackwell Companion to phonology.* Malden, MA & Oxford: Wiley-Blackwell, 1576-1596.

Kühnert, B., Nolan, F. (1999). The origin of coarticulation. In Hardcastle, W.J., Hewlett, N. (Eds.), *Coarticulation: Theory, Data and Techniques.* Cambridge: Cambridge University Press, 7-30.

Levin, J. (1987). Between epenthetic and excrescent vowels. In *Proceedings of the West Coast Conference on Formal Linguistics*, 6, 187-201.

Lyche, C. (2016). Approaching variation in PFC: the schwa level. In Sylvain, D., Durand, J., Laks, B. & Lyche, C. (Eds.), *Varieties of spoken French*. Oxford University Press, ch. 27, 352-362.

Muliačić, Z. (1973). *Fonologia della lingua italiana*. Bologna: Il Mulino.

Niebuhr, O., Lancia, L. & Meunier, C. (2008). On place assimilation in French sibilant sequences. In *Proceedings of the VII ISSP*, Strasbourg, France, 221-224.

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

Perkell, J.S. (1986). Coarticulation strategies: Preliminary implications of a detailed analysis for lower lip protrusion movements. In *Speech Communication*, 5, 47-68.

Rialland, A. (1986). Schwa et syllables en français. In *Studies in compensatory lengthening*. Dordrecht: Foris Publications, 187-226.

Ridouane, R., Fougeron, C. (2011). Schwa elements in Tashlhiyt word-initial clusters. In *Journal of LabPhon*, 2, 275-300.

Sigona, F., Stella, A., Grimaldi, M. & Gili Fivela, B. (2015). Mayday: A Software for Multimodal Articulatory Data Analysis. In *Atti del X Convegno AISV*, 173-184.

Shanin, K., Blake, S.J. (2004). A phonetic study of schwa in St'at'imcets (Lillooet Salinsh). In Gerdts, D., Matthewson, L. (Eds.), *Studies in Salinsh linguistic in honor of M. Dale Kinkade*. Missoula: University of Montana, 311-327.

Tiede, M., Shattuck-Hufnagel, S., Johnson, B., Ghohs, S., Mattheis, M., Zandipur, M. & Perkell, J. (2007). Gestural phasing in /kt/ sequences contrasting within and cross word contexts. In *Proceedings of the XVI ICPhS*, Saarbrücken, August 6-10, 521-524.