DOI: 10.17469/O2101AISV000013

STEPHAN SCHMID, SARAH WACHTER

Le ostruenti sonore nella pronuncia dell'italiano di apprendenti svizzero-tedeschi

This study examines the pronunciation of voiced obstruents in Italian as a second language of Swiss German learners. On the basis of a contrastive analysis and the predictions of current models of L2 phonology, we claim that Swiss German learners 'assimilate' the voiced obstruents of Italian to the 'lax' (but unvoiced) sounds of their L1, and that devoicing is particularly frequent in the case of /dz/, /dʒ/ and /z/+C clusters.
These hypotheses are borne out by an acoustic analysis of a corpus of read speech gathered from 10 high school students in Zurich, from which we calculated the locally unvoiced frames of 644 tokens. We conclude that L2 pronunciation is determined both by the 'perceptual filter' of the L1 and by general patterns of markedness.

1. *Introduzione*

Il presente contributo esamina la pronuncia dell'italiano da parte di un gruppo di studenti liceali zurighesi, focalizzando la realizzazione del tratto [±sonoro] nelle ostruenti, il quale rappresenta una difficoltà notoria per apprendenti con un dialetto svizzero-tedesco come L1. Dopo brevi cenni ad alcuni modelli dell'acquisizione della fonologia di una L2 (1.1) si delinea uno schizzo di fonologia contrastiva (italiano *vs*. svizzero-tedesco) relativo ai sottosistemi delle ostruenti (1.2), il che permette di formulare delle ipotesi puntuali sulle difficoltà di pronuncia degli apprendenti svizzero-tedeschi (1.3). Dopo l'illustrazione del campione (2.1), del corpus (2.2) e della metodologia di indagine (2.3) si analizza il tasso di desonorizzazione delle ostruenti sonore secondo i fattori 'parlante' (3.2), 'consonante' (3.3) e 'contesto' (3.4). La discussione dei risultati (4) permette di confermare le ipotesi circa gli effetti congiunti della L1 (che agisce come filtro percettivo) e di principi generali di marcatezza (quali il contesto fonotattico e il luogo di articolazione) sulla pronuncia di una seconda lingua.

1.1 Pronunciare i suoni di un'altra lingua: alcuni modelli teorici

Negli ultimi decenni, l'acquisizione della pronuncia in una L2 è stata oggetto di un numero via via crescente di ricerche di fonetica sperimentale, alcune delle quali fanno riferimento a modelli teorici che sono stati elaborati per rendere conto dell'emergere della struttura sonora in una seconda lingua. Per molto tempo la ricerca internazionale ha preferito l'inglese come campo d'indagine del *second language speech*, ma più recentemente è stata studiata anche la pronuncia dell'italiano da par-

te di stranieri: v. per esempio i contributi riuniti nel volume curato da Costamagna e Marotta (2008) nonché vari lavori presentati ai convegni AISV degli ultimi anni.

Non è questa la sede per passare in rassegna la ricerca in questo campo né per discutere in modo approfondito le diverse teorie che si prefiggono di spiegare la genesi dell'accento straniero (per una discussione più approfondita rimandiamo a Schmid, 2012: 633-637), ma è doveroso menzionare innanzitutto lo *Speech Learning Model* (SLM, Flege, 1995) e il *Perceptual Assimilation Model* (nella sua versione PAM-L2, Best, Tyler, 2007). Secondo lo SLM, i suoni meglio distinguibili dal punto di vista percettivo vengono anche acquisiti più facilmente (Flege, 1995: 239; ma v. già Flege, 1987). Analogamente, il PAM-L2 – sviluppato sulla base di una teoria più generale della percezione di suoni sconosciuti che Catherine Best aveva elaborato negli anni Novanta (v., ad esempio, Best, 1995) – assume che "naïve listeners" [sic] tendano ad assimilare percettivamente i suoni della L2 ai fonemi articolatoriamente più simili della loro L1 (Best, Tyler, 2007: 22; cfr. anche Avesani et al., 2008). Tutto sommato, ambedue i modelli attribuiscono quindi alla lingua materna quella funzione di 'vaglio fonologico' che era già stata individuata da Trubetzkoy (1939 [1971: 126]).

Un altro modello dell'acquisizione della fonologia di una L2 al quale occorre accennare brevemente è l'*Ontogeny Philogeny Model* (OPM, Major, 2001), che mette invece in rilievo i processi evolutivi inerenti a principi universali del linguaggio verbale umano. A questo proposito non si può non ricordare il pensiero di Roman Jakobson (come Trubetzkoy membro del Circolo di Praga e in quanto tale uno dei fondatori dello strutturalismo europeo e della fonologia moderna), e in particolare il famoso testo del 1941 (la traduzione italiana esce nel 1971), che mette in relazione lo sviluppo del linguaggio nel bambino con la perdita della parola nell'afasia, ricollegando infine questi due processi speculari a 'leggi foniche generali' che si esplicitano anche nella tipologia fonologica delle lingue del mondo.

Nell'ambito della teoria dell'acquisizione di lingue seconde, la tipologia linguistica viene chiamata in causa anche dalla *Markedness Differential Hypothesis* (MDH, Eckman, 1977 e 2008: 97-100). Questa ipotesi postula che la difficoltà delle strutture sonore della L2 dipenda non soltanto dalle differenze tra L2 e L1, ma anche dal grado di marcatezza delle strutture nelle rispettive lingue. Ad esempio, è stato dimostrato che il tratto [±sonoro] delle ostruenti di una L2 viene acquisito difficilmente nei contesti fonotattici marcati – come la posizione finale di sillaba – se la L1 non possiede ostruenti in tale contesto.

1.2 Ostruenti sonore dell'italiano e ostruenti 'leni' dello svizzero-tedesco

Iniziamo lo schizzo di fonologia contrastiva con la descrizione dell'inventario delle ostruenti nella lingua di partenza dei nostri soggetti, ovvero nel dialetto di Zurigo (cfr. Fleischer, Schmid, 2006: 144):

Figura 1 - *Inventario delle ostruenti del dialetto di Zurigo*

	bilabiali	labiodentali	alveolari	palato-alveolari	velari
occlusive	p b̥		t d̥		k g̊
fricative		f v̥	s z̥	ʃ ʒ̊	x ɣ̊
affricate		p͡f	t͡s	t͡ʃ	k͡x

Tralasciando il fonema isolato /h/ (che non compare nella tabella di fig. 1), osserviamo che questo sistema distingue due serie di ostruenti omorganiche, ovvero tre coppie di occlusive e quattro coppie di fricative. Tuttavia, come si evince dalla trascrizione fonetica con l'aggiunta di un segno diacritico ai simboli delle consonanti sonore (ad esempio [b̥]), tale opposizione non si realizza tramite la correlazione di sonorità, bensì attraverso un contrasto di 'forza' che si manifesta acusticamente più che altro come una differenza di durata (nel senso che le 'forti' durano di più delle 'leni'). Difatti, per i dialetti svizzero-tedeschi si assume tradizionalmente un'opposizione tra ostruenti 'forti' e 'leni' piuttosto che un contrasto basato sul tratto [±sonoro] (v. Fleischer, Schmid, 2006: 244-246; Nocchi, Schmid, 2006: 519-522; cfr. anche Schmid, 1999: 62). In termini articolatori e fonatori ciò implica che sia le 'forti' che le 'leni' vengono realizzate senza vibrazione alcuna delle pliche vocali.

Dal punto di vista fonologico va inoltre notato che secondo l'inventario presentato in fig. 1 la correlazione di 'tensione' non è pertinente per le affricate, dove per ognuno dei quattro luoghi di articolazione troviamo soltanto l'elemento non marcato dell'opposizione, ovvero il fonema 'forte' o [-teso]. Infine occorre precisare che nello svizzero-tedesco opera un processo di fortizione che trasforma le ostruenti 'leni' in 'semi-forti' qualora queste si trovino a formare un nesso con un'altra ostruente (Fleischer, Schmid, 2006: 248).

Passando ora alla lingua bersaglio dei nostri apprendenti, diamo un rapido sguardo all'inventario delle ostruenti dell'italiano standard (v. Bertinetto, Loporcaro, 2005: 132; cfr. anche Schmid, 1999: 133).

Figura 2 - *Inventario delle ostruenti dell'italiano standard*

	bilabiali	labiodentali	alveolari	palato-alveolari	velari
occlusive	p b		t d		k g
fricative		f v	s z	ʃ (ʒ)	
affricate			t͡s d͡z	t͡ʃ d͡ʒ	

Ciò che salta subito all'occhio in questo inventario è che il contrasto fonologico fra ostruenti omorganiche è basato sulla correlazione di sonorità, il che implica che le pliche vocali vibrino durante l'intera fase di articolazione delle ostruenti sonore. La classe delle fricative è meno numerosa in italiano rispetto allo svizzero-tedesco a causa dello statuto marginale di /ʒ/ e dell'assenza di fricative velari; mancano altresì le affricate labiali e velari. L'inventario di fig. 2 mostra inoltre che il sistema

fonologico dell'italiano sfrutta la correlazione di sonorità anche nelle caselle delle affricate alveolari e palato-alveolari (v. la presenza dei fonemi sonori /dz/ e /dʒ/), mentre nell'inventario dello svizzero-tedesco tale posizione strutturale resta vuota (cfr. fig. 1).

1.3 Ipotesi e domande di ricerca

In base alle considerazioni teoriche esposte in 1.1 e all'analisi contrastiva fornita in 1.2 possiamo quindi formulare le seguenti ipotesi e domande di ricerca circa la difficoltà che gli apprendenti svizzero-tedeschi incontrano nella pronuncia delle ostruenti dell'italiano:

- Ipotesi 1: Essendo le ostruenti 'leni' dello svizzero-tedesco simili alle ostruenti sonore italiane (di cui condividono il modo e il luogo di articolazione nonché la posizione strutturale nell'inventario fonematico, ma non la sonorità), si può presumere che gli apprendenti difficilmente riescano a realizzare le ostruenti sonore dell'italiano come tali, sostituendole invece con i foni più simili della loro L1, cioè le 'leni' (cfr. Flege, 1995; Best, Tyler, 2007).
- Ipotesi 2: Tale tendenza alla desonorizzazione delle sonore sarà più forte nel caso delle due affricate /dz/ e /dʒ/ dell'italiano, viste le rispettive lacune nell'inventario della L1; va notato però che dal punto di vista della *equivalence classification* (Flege, 1987) si tratta sempre di suoni 'simili', dato che la stessa combinazione dei due tratti di modo e luogo di articolazione esiste anche nella L1.
- Ipotesi 3: La tendenza a desonorizzare le ostruenti sonore dell'italiano sarà inoltre più forte nei gruppi consonantici formati da due ostruenti (del tipo /zb/), a causa della regola di fortizione dei dialetti svizzero-tedeschi e del carattere tipologicamente marcato di questo contesto fonotattico.

In aggiunta a queste tre ipotesi specifiche possiamo porci anche delle domande di ricerca più generali:

- Domanda di ricerca 1: Esistono altre differenze nel tasso di desonorizzazione delle ostruenti sonore – ad esempio riguardo al modo e al luogo di articolazione – che sono in linea con tendenze tipologiche (ricavabili in base alla frequenza delle consonanti nelle lingue del mondo)?
- Domanda di ricerca 2: Esistono delle differenze riguardo al tasso di desonorizzazione tra i singoli parlanti?

2. *Dati e metodo*

2.1 Il campione

Al fine di verificare le tre ipotesi specifiche e le due domande di ricerca più generali formulate nel paragrafo precedente è stato allestito un *corpus* di parlato letto prodotto da dieci studenti di un liceo di Zurigo (6 ragazze e 4 ragazzi con un'età media

di 17 anni). Al momento della registrazione, i ragazzi stavano studiando l'italiano da due anni e mezzo con tre ore di lezione alla settimana.

Tutti gli allievi usano nella comunicazione quotidiana prevalentemente il dialetto zurighese. Tuttavia, essendo la Svizzera un paese con un forte tasso di immigrazione, in molte famiglie si parla anche un'altra lingua a causa della presenza di almeno un genitore alloglotto. Fra i dieci studenti del nostro campione, solo nel caso di Joe entrambi i genitori sono svizzero-tedeschi (indichiamo i parlanti con le prime tre lettere del loro nome). Negli altri nove casi le lingue presenti in famiglia sono il cinese (Yit), il giapponese (Mik, And), il farsi (Dar), lo svedese (Eli), il ceco (Kri), due diverse varietà di tedesco della Germania (Sab) e dell'Austria (Van), e infine anche l'italiano nel caso dell'allievo San, i cui genitori provengono dalla provincia di Lecce e dal Canton Ticino.

Dal punto di vista della competenza linguistica, lo svizzero-tedesco può comunque essere considerato la varietà dominante nel repertorio linguistico di tutti i parlanti, ed è ragionevole ipotizzare che il sistema fonologico del dialetto zurighese intervenga nella loro percezione delle ostruenti italiane. Invece, vi è un solo soggetto con genitori italofoni per il quale ci si può aspettare una realizzazione delle ostruenti sonore più simile a quella di un parlante nativo.

2.2 Il *corpus*

La procedura adottata per la raccolta e l'analisi dei dati ha seguito sostanzialmente il modello di uno studio precedente sulla pronuncia del francese come L2 da parte di studenti liceali zurighesi (Schmid, 2012).

I dieci allievi sono stati registrati durante la lezione di italiano in una stanza a parte messa a disposizione dalla scuola. La registrazione è stata effettuata con un registratore digitale Fostex FR-2FLE e un microfono a cravatta Sennheiser MKE 2 P-C (direttività omnidirezionale, gamma di frequenza di 20-20.000 Hz ±23dB e coefficiente di trasmissione a vuoto di 10 mV/Pa ±2,5 dB). I file sonori sono stati salvati con una frequenza di campionamento di 44.100 Hz e una quantizzazione di 16 bit.

Durante la registrazione, ognuno dei dieci parlanti ha letto ad alta voce 19 frasi contenenti le sette ostruenti sonore dell'italiano /b d ɡ v z dz dʒ/ in sei contesti fonotattici diversi, dove le ostruenti bersaglio sono sempre seguite da vocale, mentre varia il contesto precedente:

- ##_V in posizione iniziale di enunciato;
- V#_V dopo vocale all'inizio di parola;
- C#_V dopo sonorante all'inizio di parola;
- V_V dopo vocale all'interno di parola (consonante scempia);
- V_:V dopo vocale all'interno di parola (consonante geminata);
- z_V dopo sibilante sonora.

La tabella in fig. 3 elenca le parole inserite nelle 19 frasi da leggere; le cifre prima delle parole elicitate si riferiscono alla numerazione delle frasi che vengono riprodotte in Appendice. Il *corpus* complessivo ammonta a 670 realizzazioni possibili di ostruenti sonore (10 parlanti x 67 consonanti), visto che alcune combinazioni di consonante + contesto occorrono più di una volta. Delle 670 occorrenze possibili in base allo schema riportato in fig. 3 abbiamo però dovuto scartare 26 per motivi contingenti, legati all'esecuzione del compito di lettura, per cui sono state analizzate 644 realizzazioni.

Figura 3 - *Parole elicitate secondo la consonante e il contesto*

	b	d	g	v	z	dz	dʒ
##_V	***##bV*** 1) **b**uongiorno	***##dV*** 2) **d**i 10) **d**ove	***##gV*** 3) **g**uardate 6) **G**abriele	***##vV*** 1) vorrei 14) **v**ado		***##dzV*** 18) **z**ero	***##dʒV*** 4) **g**ioca
V#_V	***V#bV*** 7) una **b**irra	***V#dV*** 1) gusto **d**i 2) di **d**ove 6) nome **d**ella 13) collega **d**i 18) a **d**ue	***V#gV*** 13) di **G**abriele	***V#vV*** 1) di **v**aniglia 11) gente **v**a		***V#dzV*** 14) allo **z**oo	***V#dʒV*** 3) è **g**ià 11) la **g**ente
C#_V	***C#bV*** 7) al **b**ar 9) il **b**ambino 12) il **b**ambino	***C#dV*** 15) il **d**ialogo 16) un **d**ubbio 19) il **d**isgelo	***C#gV*** 1) al **g**usto	***C#vV*** 4) il **V**erona 17) in **V**altellina		***C#dzV*** 13) in **z**ona	***C#dʒV*** 1) un **g**elato
V_V	***VbV*** 10) a**b**itate 13) a**b**ita	***VdV*** 9) sgri**d**a 11) sta**d**io 14) va**d**o 16) enciclope**d**ia	***VgV*** 13) ccl**g**a 15) dia**g**o	***VvV*** 2) do**v**e 10) do**v**e 19) primave**r**a	***VzV*** 6) ro**s**a		***VdʒV*** 3) mezzo**g**iorno 14) cu**g**ino
V_:V	***Vb:V*** 16) a**bb**iamo 16) du**bb**io	***Vd:V*** 12) contra**dd**ice 17) A**dd**a	***Vg:V*** 15) le**gg**o	***Vv:V*** 5) a**vv**ersari		***V#d:zV*** 3) mezzogiomo 8) realizzare	***Vd:ʒV*** 4) o**gg**i 6) le**gg**e 16) le**gg**iamo
z_V	***zb*** 8) **sb**agliato	***zd*** 15) **sd**raio	***zg*** 9) **sg**rida	***zv*** 2) **sv**izzero			***zdʒ*** 19) di**sg**elo

2.3 Procedura di analisi

L'analisi acustica delle registrazioni sonore è stata effettuata mediante il programma *Praat* (Boersma, Weenink, 2015). In un primo passo sono state segmentate e annotate su un *TextGrid* le realizzazioni delle ostruenti in questione, per calcolarne successivamente la percentuale di desonorizzazione (ovvero la percentuale di segnale non periodico in quella sezione temporale del segnale); in *Praat*, la funzione *fraction of locally unvoiced frames* viene fornita dal *voice report* che si trova nella finestra *edit* sotto il menu *pulses*.

L'analisi statistica dei dati è stata effettuata mediante i programmi *Microsoft Excel* e SPSS.

3. *Risultati*

3.1 Il tasso globale di desonorizzazione

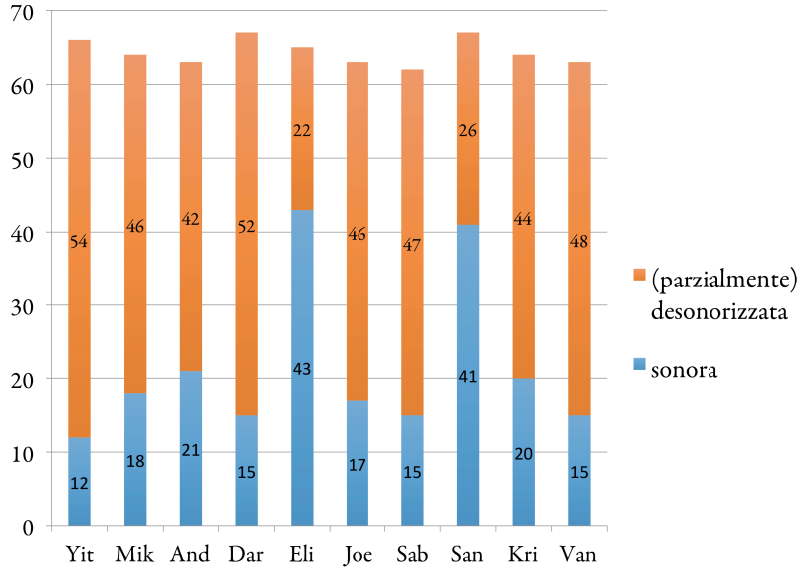
Delle 644 occorrenze effettivamente prese in considerazione, 217 sono state realizzate come pienamente sonore, mentre solo 12 sono state completamente desonorizzate. Invece, in ben 415 realizzazioni (ovvero il 64.4% dei casi) abbiamo incontrato delle soluzioni 'intermedie', con un tasso di desonorizzazione che va dall'11% fino al 97%. Considerando il numero esiguo delle ostruenti sonore completamente desonorizzate, d'ora in poi le considereremo assieme alle realizzazioni parzialmente desonorizzate (ottenendo quindi solo due categorie: 'sonore' e 'più o meno desonorizzate'). In effetti, se un allievo produce un fono [d] desonorizzato all'86%, possiamo dedurre che nella fonologia dell'interlingua tale fono corrisponda a un fonema /d̥/ e che la breve fase periodica all'inizio della fase di occlusione corrisponda – almeno nel caso che l'occlusiva venga preceduta da una vocale – a un fenomeno di coarticolazione, cioè alla cosiddetta *VoiceOFFsetTime*.

Come primo risultato globale possiamo comunque constatare che la realizzazione di ostruenti completamente sonore crea non pochi problemi agli apprendenti svizzero-tedeschi, dato che in quasi due terzi dei casi essi non riescono a produrle come tali. I dati mostrano però anche una notevole variazione secondo i fattori 'parlante', 'consonante' e 'contesto' che analizzeremo separatamente qui di seguito.

3.2 Il fattore 'parlante'

Cerchiamo innanzitutto di rispondere alla domanda di ricerca 2 formulata in 1.3: esistono delle differenze riguardo al tasso di desonorizzazione tra i singoli parlanti?

Figura 4 - *Numero delle realizzazioni di ostruenti sonore secondo i parlanti*



La fig. 4 mostra il grado di desonorizzazione delle ostruenti sonore per ognuno dei dieci parlanti, indicando il numero di realizzazioni nelle due categorie 'ostruente sonora' e 'ostruente (parzialmente o totalmente) desonorizzata'.

Tutto sommato possiamo dire che il quadro è piuttosto omogeneo, nel senso che in otto parlanti predominano le realizzazioni desonorizzate. Fanno però eccezione i due parlanti Eli e San, in cui le ostruenti completamente sonore (43 e 41 realizzazioni rispettivamente) superano quelle (parzialmente) desonorizzate (22 e 26 realizzazioni rispettivamente).

Per quanto riguarda la durata della fase aperiodica nelle singole realizzazioni desonorizzate (quelle all'interno della porzione arancione delle barre di fig. 4), la gamma delle percentuali varia notevolmente per tutti i parlanti; ad esempio, la parlante Van produce da un lato una [dz] desonorizzata al 92% e dall'altro una [b] desonorizzata al 14%. Considerando gli otto parlanti con un comportamento relativamente omogeneo (tutti meno Eli e San), la percentuale media di desonorizzazione calcolata per la durata delle ostruenti parzialmente sonore è del 47%, con una gamma che va dal 37% nel parlante And al 57% nel caso della parlante Yit e una deviazione standard del 7%.

Da questo quadro generale si distanziano però di nuovo i parlanti Eli e San che presentano un grado medio di desonorizzazione del solo 18%. Tenendo conto dei repertori linguistici parlati in famiglia (cfr. 2.1), possiamo forse spiegare il comportamento di San con l'italofonia dei suoi genitori (in un certo senso questo parlante può essere considerato bilingue). Nella famiglia di Eli si parla anche lo svedese, una lingua nella quale esistono ostruenti sonore a tutti gli effetti (cfr. Engstrand, 1999).

Considerando ora soltanto gli altri otto parlanti, troviamo comunque una certa correlazione tra il numero delle realizzazioni desonorizzate prodotte dai singoli parlanti e le durate della fase aperiodica all'interno di questi foni ($r = 0.76$, $R^2 = 0.57$): in altre parole, nei parlanti che pronunciano un numero più alto di ostruenti (parzialmente) desonorizzate, la fase aperiodica all'interno delle realizzazioni (parzialmente) desonorizzate è mediamente anche più lunga. Per l'analisi statistica delle realizzazioni desonorizzate abbiamo però dovuto rinunciare all'applicazione di un test parametrico (in questo caso di un'ANOVA univariata) a causa dei risultati di due test preliminari. Infatti, secondo il test di Shapiro-Wilk i valori sono distribuiti normalmente solo per 5 parlanti su dieci (Mik, And, Eli, Sab e Kri). Inoltre, in base al test di Levene non vi è uguaglianza della varianza tra i dieci parlanti ($p<.001$). Comunque, il test (non-parametrico) di Kruskal-Wallis indica che tra i dieci parlanti esistono delle differenze significative riguardo al grado di desonorizzazione delle ostruenti sonore ($\chi^2(9)=93{,}584$, $p<.001$).

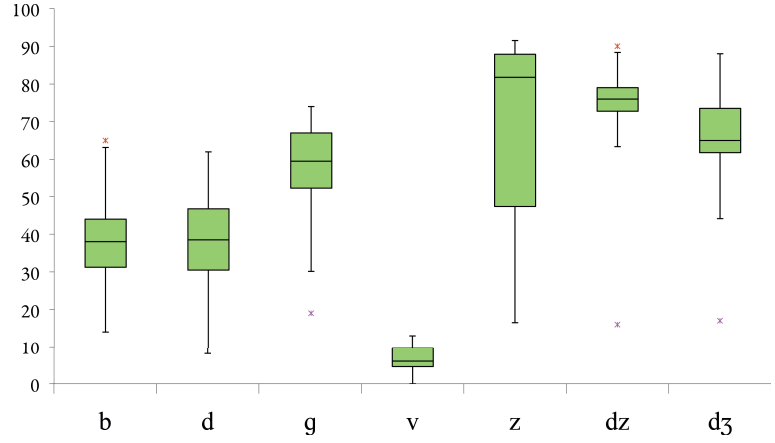
3.3 Il fattore 'consonante'

Nemmeno nel caso del fattore 'consonante' i test preliminari hanno permesso di applicare un test parametrico. Anche se consideriamo unicamente le realizzazioni (parzialmente) desonorizzate (cfr. fig. 4), il test di Shapiro-Wilk mostra una distribuzione normale soltanto per le tre ostruenti /v/, /ɡ/ e /dz/, ma non per /b/, /d/,

/z/ e /dʒ/. Inoltre, secondo il test di Levene le percentuali di desonorizzazioni non hanno la stessa varianza nelle sette ostruenti. Ricorrendo invece al test non parametrico di Kruskal-Wallis, si ottiene senz'altro un effetto significativo del fattore 'consonante' sulle percentuali di desonorizzazione ($\chi^2(6)=266{,}048$, $p<.001$).

Tale effetto è visibile anche a livello descrittivo nella fig. 5, i cui boxplot illustrano – per le sette ostruenti sonore prodotte da tutti e dieci i parlanti – la dispersione delle percentuali di desonorizzazione da 0% a 100% (comprese quindi le realizzazioni completamente sonore e sorde). I boxplot sono in effetti interpretabili in base alle ipotesi formulate nel paragrafo 1.3. Iniziando le nostre considerazioni con le tre occlusive /b d g/, notiamo ad esempio un tasso di desonorizzazione molto simile per le bilabiali e le alveolari (con una media del 37%), mentre le velari sono colpite più frequentemente dalla desonorizzazione (con una media del 56%).

Figura 5 - *Percentuale media della desonorizzazione secondo il fattore 'consonante'*

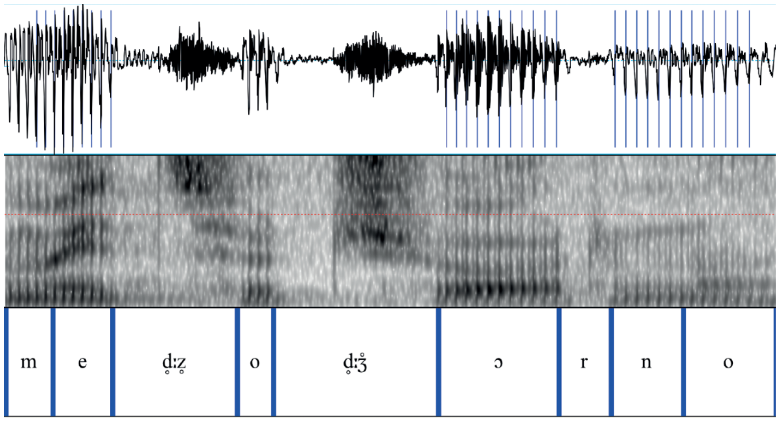


La maggiore desonorizzazione delle velari non può essere determinata dalla struttura della L1, per cui è più promettente cercarne la causa – sulla scia della più generale domanda di ricerca 2 – su eventuali tendenze universali nella distribuzione della consonante /g/. Ricordiamo, tra parentesi, che già Jakobson (1941 [1971]: 89-90) aveva postulato la precedenza delle consonanti labiali e dentali nell'acquisizione del linguaggio nei bambini. Ebbene, dagli studi di tipologia fonologica emerge che tra le occlusive sonore (di per sé marcate rispetto alle sorde) è più probabile che in un inventario fonematico manchi /g/ che non /b/ e /*d/ (Maddieson, 1984: 36). Questo risultato viene confermato da una consultazione della banca dati PHOIBLE allestita recentemente (Moran et al., 2014): le occlusive labiali (/b/) e coronali (/d/ e /d̪/) sono presenti rispettivamente nel 71% e 72% delle 2155 lingue documentate, laddove /g/ si trova solo nel 64% dei casi. Una possibile causa aerodinamica di questa differenza viene indicata da Ohala (1992: 8) nella difficoltà di mantenere la fonazione con un condotto vocale bloccato e piuttosto ristretto come nel caso del luogo di articolazione velare.

Tornando alla fig. 5 e volgendo lo sguardo verso il lato destro del grafico si nota un tasso di desonorizzazione molto elevato per i due fonemi /dz/ e /dʒ/. Per illu-

strare tale fenomeno riportiamo nella fig. 6 l'oscillogramma e lo spettrogramma della parola *mezzogiorno* pronunciata dal parlante Dar: la desonorizzazione delle due affricate viene evidenziata anche dall'assenza di strisce verticali blu che caratterizza le parti aperiodiche nel segnale acustico.

Figura 6 - *Oscillogramma e spettrogramma della parola* mezzogiorno *pronunciata dal parlante Dar*



Constatiamo a questo punto che la forte tendenza alla desonorizzazione delle due affricate sonore è conforme all'ipotesi 2 (v. 1.3). Oltre alla mancanza di questa posizione strutturale nell'inventario fonematico della L1 (ovvero all'assenza di affricate leni nello svizzero-tedesco), la difficoltà di pronunciare un'affricata sonora sembra però essere dovuta anche a una complessità articolatoria intrinseca di questo tipo di consonante. Sulla scia di Jakobson (1941 [1971]: 56, 52) possiamo infatti assumere che il modo di articolazione delle affricate sia marcato rispetto a quello delle fricative, le quali sarebbero a loro volta marcate rispetto alle occlusive.

Guardando infine la parte centrale della fig. 5, troviamo un comportamento molto diverso tra le due fricative, con soltanto 6% di desonorizzazioni della labiodentale rispetto a 68% nel caso della sibilante. Questo scarto non trova una spiegazione in considerazioni tipologiche (dato che /z/ e /v/ compaiono rispettivamente nel 31% e 29% delle lingue documentate nella banca dati PHOIBLE), ma sarà da ricondurre piuttosto all'azione della prima lingua degli apprendenti. Per quanto riguarda /v/, la facilità con cui essi riescono a produrre un suono labiodentale sonoro si spiega in effetti con la presenza dell'approssimante /ʋ/ nei dialetti svizzero-tedeschi (cfr. Fleischer, Schmid, 2006: 244; Nocchi, Schmid, 2006: 522-523). Invece, la difficoltà di /z/ non deriva tanto da proprietà 'paradigmatiche' di questa consonante quanto da un processo 'sintagmatico' che vedremo nel prossimo paragrafo.

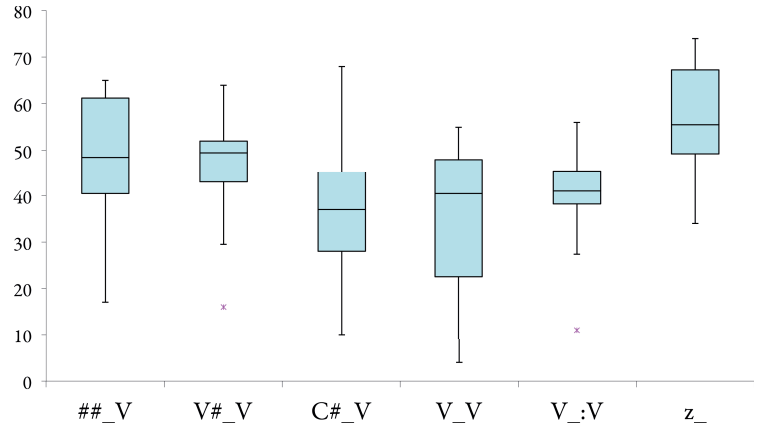
3.4 Il fattore 'contesto'

Anche per l'ultimo dei tre fattori presi in considerazione – il contesto fono(sin)tattico – abbiamo dovuto escludere l'applicazione di un test parametrico in base

ai risultati dei due test preliminari. Infatti, secondo il test di Shapiro-Wilk i valori sono distribuiti normalmente soltanto in due contesti, cioè all'inizio dell'enunciato (contesto ##_V) e all'inizio di parola dopo sonorante (contesto C#_V), ma non negli altri quattro contesti (V#_V, V_V, V_:V, z_); inoltre, il test di Levene non ha fornito una varianza uguale delle percentuali di desonorizzazione. Di conseguenza abbiamo optato anche in questo caso per un test non parametrico, il Kruskal-Wallis per l'appunto, che indica comunque un effetto significativo anche per il fattore 'contesto' ($\chi^2(5)=76{,}027$, $p<.001$).

Passando quindi alla statistica descrittiva, vediamo in fig. 7 dei boxplot che illustrano il diverso grado di desonorizzazione delle ostruenti sonore secondo la loro posizione nella parola e nell'enunciato, distinguendo i sei contesti fonotattici descritti in 2.2 per i dati dell'intero *corpus*:

Figura 7 - *Percentuale di desonorizzazione secondo il fattore 'contesto fonotattico'*



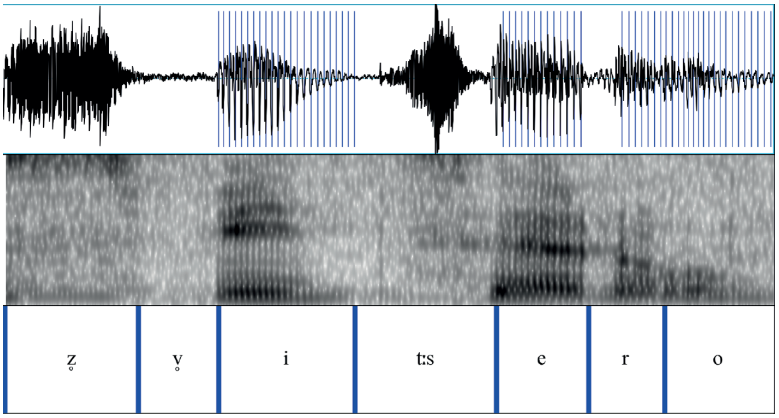
Dai boxplot si evince che la desonorizzazione è più contenuta nel contesto intervocalico all'interno di parola, il che viene confermato dalle medie che ammontano al 35% per le scempie (V_V) e al 38% per le geminate (V_:V). Questo risultato è in linea con le predizioni della *Markedness Differential Hypothesis* e può essere spiegato con la relativa facilità fisiologica di mantenere la fonazione quando l'ostruente è preceduta e seguita da una vocale. Un simile tasso di desonorizzazione, con una media del 38%, si trova al confine iniziale della parola quando precede una sonorante (C#_V), mentre sono più alti i valori medi dopo vocale (V#_V) e quando la consonante si trova in posizione iniziale di enunciato (##_V): le rispettive medie ammontano al 45% e al 48%. Il valore leggermente più alto risulta alquanto plausibile per la posizione iniziale di enunciato, ma non tanto per la posizione iniziale di parola.

Risulta invece chiaro che la desonorizzazione compare con maggiore incidenza quando l'ostruente sonora è preceduta da una sibilante (z_). In questo contesto troviamo una media di desonorizzazione del 57% che va ricollegata al 68% riscontrato per /z/ nel paragrafo precedente. È evidente come nel nesso /z/ + ostruente sonora

operi il processo di fortizione dello svizzero-tedesco descritto in 2.1, che favorisce notevolmente la pronuncia sorda di ambedue le consonanti nella L2.

Per illustrare questo fenomeno riportiamo infine l'oscillogramma e lo spettrogramma della parola *svizzero* pronunciata dalla parlante Mik:

Figura 8 - *Oscillogramma e spettrogramma della parola* svizzero *pronunciata dalla parlante Mik*



Anche in questo caso l'assenza di strisce verticali rivela la totale aperiodicità del segnale acustico durante l'articolazione dei primi due segmenti della parola, rivelando la completa desonorizzazione del nesso /zv/.

4. *Discussione e conclusioni*

Avviandoci a qualche riflessione interpretativa sui risultati appena esposti, ci pare opportuno addurre prima alcuni elementi di confronto con lo studio precedente sulla pronuncia delle ostruenti francesi da parte di un'altra classe di studenti liceali zurighesi (cfr. Schmid, 2012). Innanzitutto va rilevata una differenza sostanziale riguardo alla distribuzione delle tre categorie nel *corpus* di francese L2: su 340 occorrenze, 106 (31.2%) sono completamente sonore e 62 (18.2%) parzialmente desonorizzate, mentre ben 172 (50.6%, in pratica la metà) sono completamente desonorizzate. Invece, delle 644 occorrenze analizzate nel nostro corpus di italiano L2, 217 (33.7%) vengono realizzate come sonore e 415 (64.4%) come parzialmente desonorizzate, mentre vi sono soltanto 12 occorrenze (1.9%) completamente sorde (cfr. 3.1). Tale differenza è in parte dovuta al fatto che la lingua francese ammette ostruenti sonore anche in posizione finale di parola e quindi in un contesto particolarmente suscettibile di innescare il processo di desonorizzazione (cfr. Eckman, 2008). Tuttavia, nel *corpus* di francese L2 le realizzazioni desonorizzate occorrono anche in altri contesti fonotattici (cfr. Schmid, 2012: 645), per cui possiamo affermare che il campione analizzato nel presente studio riesce ad avvicinarsi di più alla struttura sonora della lingua bersaglio.

È difficile spiegare questa differenza nel comportamento delle due classi liceali (sul quale possono influire molteplici fattori quali la predisposizione all'acquisizione di una L2 dei singoli allievi, la qualità dell'insegnamento, ecc.), ma si può forse considerare il fatto che il francese rappresenta una materia scolastica obbligatoria (spesso non molto amata), laddove l'italiano è una lingua opzionale per la quale gli allievi potrebbero essere più motivati. È vero che nel nostro *corpus* quasi due terzi delle ostruenti sono almeno in parte desonorizzate, ma forse vale la pena insistere sia sul carattere parziale di questa desonorizzazione sia sulla presenza – in tutti i locutori – di ostruenti completamente sonore. È proprio questa forte variabilità riscontrata nei dati a portarci a una visione non tanto drammatica del livello raggiunto dagli allievi nella pronuncia dell'italiano. In base al carattere meno stabile delle interlingue in quanto sistemi in via di sviluppo, la produzione in una lingua seconda presenta spesso una maggiore variabilità rispetto a quella nella lingua nativa; in altre parole, si tratta di un tipo di variabilità che denota un potenziale di avvicinamento alla norma della lingua bersaglio.

Ciononostante non possiamo negare che la pronuncia delle ostruenti sonore costituisce una difficoltà notevole per gli apprendenti zurighesi; forse si tratta addirittura di una caratteristica saliente di ciò che potremmo definire come 'accento svizzero' nella pronuncia dell'italiano.

Tornando quindi al punto di partenza delle nostre riflessioni, i risultati di questo studio permettono di confermare in sostanza le tre ipotesi che avevamo formulato sulla base dell'analisi contrastiva tra L1 e L2 nonché sulla base dei postulati dei principali modelli teorici della fonologia delle interlingue:

- Ipotesi 1: In base alla somiglianza percettiva e all'equivalenza strutturale con le consonanti 'leni' della L1, le ostruenti sonore dell'italiano vengono in buona parte 'assimilate' alla rispettiva categoria dello svizzero-tedesco. Questi risultati sono in linea con le predizioni dell'SLM e del PAM-L2.
- Ipotesi 2: Tra le sette ostruenti sonore dell'italiano, le due affricate sono maggiormente soggette alla desonorizzazione, date l'assenza di un equivalente strutturale nella L1 e la complessità intrinseca di questi foni.
- Ipotesi 3: Un tasso di desonorizzazione più alto si trova anche nei nessi di sibilante + ostruente sonora come riflesso di un processo fonotattico della L1.

Per quanto riguarda la seconda domanda di ricerca più generale – le differenze tra i parlanti – abbiamo constatato un numero più alto di realizzazioni sonore presso due allievi, di cui uno è di origine italiana. Ai fini del nostro studio è però più rilevante la prima domanda di ricerca: l'eventuale coincidenza tra alcuni risultati nel *corpus* con tendenze note dalla tipologia fonologica. Possiamo interpretare in questo senso da un lato la minore desonorizzazione delle ostruenti intervocaliche e dall'altro la maggiore desonorizzazione delle occlusive velari rispetto agli altri luoghi di articolazione. Concludiamo quindi che il fenomeno della desonorizzazione nella pronuncia delle ostruenti sonore dell'italiano è dovuto sia all'influsso del sistema fonologico dello svizzero-tedesco sia a principi generali della struttura sonora del linguaggio.

Ringraziamenti

Siamo grati a Letizia Könz, insegnante di italiano nel *Realgymnasium Rämibühl* di Zurigo, nonché agli allievi della classe 5i per la loro partecipazione a questa ricerca. Grazie a Sandra Schwab per la preziosa consulenza statistica. Ringraziamo infine tre revisori anonimi per le osservazioni critiche molto dettagliate sulla prima versione del testo; ovviamente, i due autori rimangono gli unici responsabili di eventuali errori o imprecisioni.

Bibliografia

Avesani, C., Vayra, M., Best, C. & Bohn, O. (2008). Fonologia e acquisizione. In che modo l'esperienza della lingua materna plasma la percezione dei suoni del linguaggio. In Costamagna, L., Marotta, G. (Eds.), 15-41.

Bertinetto, P.M., Loporcaro, M. (2005). The sound pattern of Standard Italian, as compared with the varieties spoken in Florence, Milan and Rome. In *Journal of the International Phonetic Association*, 35 (2), 131-151.

Best, C. (1995). A direct realist view of cross-language speech perception. In Strange, W. (Ed.), 171-204.

Best, C., Tyler, M. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In Bohn, O., Munro, M. (Eds.), *Language experience in second language speech learning: In honor of James Emil Flege.* Amsterdam: Benjamins, 13-33.

Boersma, P., Weenink, D. (2015). *Praat: doing phonetics by computer* [Computer program]. Version 5.4.09, retrieved 1 June 2015 from http://www.praat.org/.

Costamagna, L., Marotta, G. (Eds.) (2008). *Processi fonetici e categorie fonologiche nell'acquisizione dell'italiano*. Pisa: Pacini.

Eckman, F. (1977). Markedness and the Contrastive Analysis Hypothesis. In *Language Learning*, 27, 315-330.

Eckman, F. (2008). Typological markedness and second language phonology. In Hansen Edwards, J., Zampini, M. (Eds.), *Phonology and second language acquisition.* Amsterdam: Benjamins, 95-115.

Engstrand, O. (1999). Swedish. In *Handbook of the International Phonetic Association.* Cambridge: Cambridge University Press, 140-142.

Flege, J. (1987). The production of "new" and "similar" phones in a foreign language: evidence for the effect of equivalence classification. In *Journal of Phonetics*, 15, 47-65.

Flege, J. (1995). Second language speech learning. Theory, findings, and problems. In Strange, W. (Ed.), 233-277.

Fleischer, J., Schmid, S. (2006). Zurich German. In *Journal of the International Phonetic Association*, 36 (2), 243-255.

Jakobson, R. (1941). Kindersprache, Aphasie und allgemeine Lautgesetze. In *Uppsala Universitets Årsskrift*, 1-83 [trad. it. 1971, *Il farsi e il disfarsi del linguaggio. Linguaggio infantile e afasia.* Torino: Einaudi].

Maddieson, J. (1984). *Patterns of sound*. Cambridge: Cambridge University Press.

Major, R. (2001). *Foreign accent. The ontogeny and philogeny of second language acquisition.* Mahwah, NJ: Lawrence Erlbaum Associates.

Moran, S., McCloy, D. & Wright, R. (2014). *PHOIBLE Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology (Available online at http://phoible.org, Accessed 4.06.15).

Nocchi, N., Schmid, S. (2006). Le consonanti labiodentali dello svizzero tedesco. In Savy, R., Crocco, C. (Eds.), *Analisi prosodica. Teorie, modelli e sistemi di annotazione* (Atti del 2° Convegno Nazionale AISV. Salerno, 30 Novembre-2 Dicembre 2005). Torriana: EDK Editore, 518-537.

Ohala, J. (1992). What's cognitive, what's not, in sound change. In Kellermann, G., Morrissey, M. (Eds.), *Diachrony within synchrony: language history and cognition*. Frankurt a.M.: Peter Lang, 309-355.

Schmid, S. (1999). *Fonetica e fonologia dell'italiano.* Torino: Paravia.

Schmid, S. (2012). The pronunciation of voiced obstruents in L2 French: a preliminary study of Swiss German learners. In *Poznań Studies in Contemporary Linguistics*, 48 (4), 627-659.

Strange, W. (Ed.) (1995). *Speech perception and linguistic experience. Issues in cross-language research*. Timonium, MD: York Press.

Trubetzkoy, N. (1939). *Grundzüge der Phonologie.* Göttingen, Vandenhoeck & Ruprecht [trad. it. 1971, *Fondamenti di fonologia.* Torino: Einaudi].

Appendice

Le 19 frasi del questionario:

1. Buongiorno. Vorrei un gelato al gusto di vaniglia.
2. Di dove sei? – Sono svizzero.
3. Guardate, è già mezzogiorno.
4. Gioca il Verona oggi?
5. Gli avversari si abbracciano.
6. Gabriele legge il libro "*Il nome della rosa*".
7. Prendiamo una birra al bar.
8. È sbagliato non realizzare i suoi sogni.
9. La mamma sgrida il bambino.
10. Dove abitate?
11. La gente va allo stadio.
12. Il bambino contraddice il padre.
13. Il collega di Gabriele abita in zona.
14. Vado allo zoo con mio cugino.
15. Mi sdraio al sole e leggo il dialogo.
16. Se abbiamo un dubbio leggiamo l'enciclopedia.
17. L'Adda scorre in Valtellina.
18. Zero a due, questo è il risultato.
19. A primavera c'è il disgelo.