

# L'ACQUISIZIONE DEL SISTEMA VOCALICO ITALIANO: UNO STUDIO ACUSTICO SU PARLANTI GERMANOFONI E ROMENOFONI

*Claudia Roberta Combei<sup>1</sup>, Ottavia Tordini<sup>2</sup>, Fabio Ardolino<sup>3</sup>*

## 1. INTRODUZIONE

Nel lungo e complicato percorso che porta idealmente al raggiungimento della totale competenza linguistica in una lingua straniera (L2<sup>4</sup>) alcuni ostacoli sono più difficili da superare di altri. In particolare, la competenza fonologica, che nella lingua madre (L1) viene pienamente padroneggiata nei primi anni dell'età scolare, risulta particolarmente impegnativa da apprendere nella L2 (Combei, Marotta, 2019). Questo è dovuto sostanzialmente al complesso processo di “risintonizzazione” dei meccanismi di percezione e produzione (Combei, 2019), la cui incompiutezza si manifesta anche nella presenza di tratti segmentali e soprasegmentali della L1 nella L2. Queste discrepanze vengono spesso percepite dall'orecchio dell'ascoltatore nativo: un fenomeno comunemente chiamato “accento straniero”<sup>5</sup>.

Negli ultimi decenni l'acquisizione dei tratti segmentali – in uguale misura vocalici e consonantici – è stata ampiamente indagata negli studi sulla L2, soprattutto per quanto riguarda l'inglese come lingua *target* (Piske *et al.*, 2011; Lin, 2013; Sereno, 2016). Tuttavia, diversi studi hanno esaminato anche i correlati acustici dell'italiano L2. Costamagna (2008), ad esempio, basandosi sul framework teorico dell'*Ontogeny Philogeny Model* (Major, 2001), ha condotto uno studio longitudinale sulla fonologia dell'interlingua di un parlante brasiliano. I risultati mostrano la frequente sostituzione delle affricate con le fricative, a causa della minore marcatezza di queste ultime. L'autrice conclude affermando che la lunga e complessa acquisizione delle affricate italiane da parte del parlante brasiliano conferma come la similarità e la marcatezza abbiano un forte effetto sulla *performance*. Per quanto riguarda le vocali, in uno studio sulla misurazione dell'accento straniero, Pellegrino (2012) afferma che la durata delle vocali toniche e atone, insieme ad altre caratteristiche segmentali e soprasegmentali, siano dei validi indicatori per distinguere chiaramente gli apprendenti cinesi analizzati dai parlanti madrelingua italiani.

A questo proposito, notiamo che l'acquisizione del sistema vocalico italiano *in toto* è stata finora oggetto di pochi studi nell'ambito della fonetica acustica. De Meo, Vitale, Xu

<sup>1</sup> Università di Bologna.

<sup>2</sup> Università di Pisa.

<sup>3</sup> Università di Pisa, Aix-Marseille Université.

L'articolo è il risultato della collaborazione tra i tre autori. La responsabilità del lavoro viene così attribuita: § 1, § 2.1, § 3.3, § 3.4, § 4: C.R. Combei; § 2.2, § 2.4, § 3.2: O. Tordini; § 2.3, § 3.1: F. Ardolino.

<sup>4</sup> Usiamo L2 come concetto generico per riferirci a tutte le lingue imparate successivamente alla lingua madre.

<sup>5</sup> Una definizione generalmente accettata dell'accento straniero è stata formulata da Munro: «Non-pathological speech produced by L2 learners that differs in partially systematic ways from the speech characteristics of native speakers» (Munro, 1998: 139).

(2012) hanno analizzato acusticamente le vocali di sei apprendenti cinesi di italiano L2, estratte da un *corpus* di parlato letto. I risultati mostrano come, a prescindere dal livello di competenza in italiano, gli apprendenti non riescano a distinguere tra vocali semi-aperte e semi-chiuse e tra vocali toniche e atone. Inoltre, le autrici affermano che solo i fonemi, e non gli allofoni, sono trasferiti dal cinese all'italiano.

Romito e Frontera (2015) hanno invece analizzato le opposizioni /e – i/ e /o – u/ nelle vocali atone in sillaba finale, un'opposizione che gli apprendenti galiziani analizzati riducono a /e – o/ nella loro L1. Tuttavia, i risultati mostrano che questi parlanti siano in grado di distinguere le opposizioni in italiano.

## 2. LO STUDIO

### 2.1. Obiettivi

Il presente contributo si concentra sul vocalismo tonico di dodici apprendenti adulti di italiano L2, rispettivamente 6 germanofoni e 6 romenofoni, paragonati a un numero congruo di soggetti madrelingua italiani (v. § 3.1 per la descrizione del *corpus*). Questo tipo di confronto può considerarsi assente, fino ad oggi, nel panorama degli studi di fonetica sperimentale sull'italiano L2. Nello specifico, il nostro lavoro intende valutare se i parlanti romeni e tedeschi siano in grado di produrre accuratamente le vocali italiane e di distinguere l'opposizione tra le vocali medio-aperte e medio-chiuse. Tale aspetto è meritevole di indagine in quanto i romenofoni non possiedono il valore di opposizione per l'apertura vocalica (/ɔ/ vs. /o/ e /ɛ/ vs. /e/), mentre i tedescofoni presentano valori di opposizione per la durata vocalica, nonché un inventario più ricco rispetto all'italiano. (v. § 2.4 per una presentazione dei sistemi vocalici a confronto).

Al fine di rispondere allo scopo della ricerca si testeranno le ipotesi dello *Speech Learning Model* (Flege, 1995, 2007), un modello percettivo costruito principalmente intorno al ruolo della L1 e dell'*input* nell'acquisizione della fonologia della L2, due elementi studiati ampiamente da Flege (2009). Utilizzeremo parallelamente come basi teoriche anche le teorie dell'*input* proposte da Krashen (1981, 1985). Ci avvarremo inoltre di analisi strumentali (estrazioni di valori formantici e durata), statistica descrittiva e inferenziale (regressione lineare) e tecniche di *clustering* (*k-Means*) per comprendere e misurare il ruolo delle variabili sociolinguistiche nell'apprendimento del vocalismo italiano e per ipotizzare possibili implicazioni nell'insegnamento dell'italiano a stranieri.

### 2.2. Framework teorico: lo *Speech Learning Model*

Uno dei modelli teorici di riferimento per l'acquisizione di suoni L2 in apprendenti adulti è indubbiamente lo *Speech Learning Model* (SLM: Flege, 1995, 2007). Il *framework* si focalizza sulla produzione dei suoni non nativi: esso considera la distanza *cross-linguistica* tra suoni L1 e L2 – a livello di proprietà acustiche – come predittore della relativa difficoltà o facilità nella percezione e nella produzione del parlato non nativo (Bergmann *et al.*, 2016). Alla base dello SLM vi è l'assunto per cui i sistemi fonetici L1 e L2 siano “contenuti” in uno spazio comune e siano distinti su base allofonica (e non fonemica) (Chang, 2012), lungo un *continuum* di “poli” che si estende da “identico” a “nuovo”, attraverso uno stadio intermedio definito “simile” (Vayra, Avesani, Best, Bohn, 2012: 122-123).

In base a questi principi, possono verificarsi due processi di interazione: dissimilazione o assimilazione. Una sufficiente dissimilarità tra il nuovo suono L2 e la categoria L1 più vicina innesca la formazione di una nuova categoria fonetica, che permette al fono non nativo di essere prodotto accuratamente. Ciò significa che i foni L2 che sono sufficientemente divergenti da qualsiasi fono L1 subiranno una “dissimilazione” (o polarizzazione) rispetto al fono L1 più vicino e saranno quindi percepiti e prodotti accuratamente dagli apprendenti di L2. Di conseguenza, le proprietà fonetiche dei suoni in questione divergeranno l’una dall’altra. Il meccanismo di “assimilazione” si attiva invece nel caso in cui un suono L2 sia percepito come “identico” o “simile” a un suono esistente nel sistema nativo (cioè un esempio più o meno deviante di una categoria fonetica L1). Il processo che ostacola la formazione di una nuova categoria è chiamato da Flege “*equivalence classification*”. In questo caso, «a single phonetic category will be used to process perceptually linked L1 and L2 sounds» (Flege, 1995: 239)<sup>6</sup>.

Gli apprendenti di L2 saranno quindi in grado di percepire le proprietà fonetiche dei suoni L2 in modo più o meno accurato a seconda dei vari scenari qui presentati, anche ben oltre la cosiddetta “soglia critica” di acquisizione bilingue (Flege, 2007). Per tali motivi, il framework si rivela particolarmente adatto ai fini della presente ricerca; di conseguenza, riteniamo che la descrizione del sistema linguistico nativo sia degli apprendenti L2, sia del gruppo di controllo, sia essenziale per misurare e interpretare la qualità della pronuncia dei suoni non-nativi. Nello specifico, in questo lavoro si valuterà sperimentalmente se, rispetto all’inventario vocalico L1, l’assenza/presenza di categorie fonetiche in italiano L2 possa predire difficoltà o facilità nella percezione e produzione delle vocali non native.

### 2.3. Il ruolo dell’input

Uno studio che intenda valutare un qualsiasi aspetto dell’acquisizione di una L2 non può non confrontarsi con il quadro teorico dell’*input hypothesis*, per come sviluppata in più fasi successive da Stephen Krashen (1981, 1985).

Nei modelli derivati da tale teoria (*input-related models*) un ruolo cruciale è assunto, nella costruzione di una competenza per una L2, dall’*input*, inteso come qualsiasi segmento percepibile prodotto nella lingua in acquisizione da un interlocutore. Affinché tale segmento cooperi alla progressione della competenza, tuttavia, è necessario che presenti specifiche caratteristiche di comprensibilità (da cui la nozione di “*input comprensibile*”, in inglese *comprehensible input*, CI).

Un *input* in ingresso assume caratteristiche di comprensibilità se non prodotto in condizioni di rumore che ne ostacolano la percezione, se perviene al soggetto ricevente in idonee condizioni di mobilitazione attenzionale e motivazionale e, soprattutto, se contiene informazioni linguistiche in grado di innestarsi in modo coerente su quelle già contenute nella competenza L2 dell’apprendente (per usare la schematizzazione di Krashen, un CI è un *input* che ha struttura  $i+1$ , laddove  $i$  è la competenza posseduta dall’apprendente fino al momento dell’ingresso dell’*input*).

In sintesi, l’adozione, anche parziale, di un quadro di riferimento di tipo *input-related* pone qualsiasi indagine relativa all’acquisizione di una L2 di fronte a una serie di punti non trascurabili:

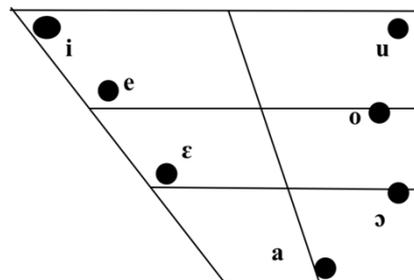
<sup>6</sup> «This is supposed to lead to long-term pronunciation problems, since similar phones are less easily perceived» (de Leeuw, 2008: 29).

- 1) esiste una cesura netta fra *language acquisition*, acquisizione spontanea del linguaggio e *language learning*, apprendimento del linguaggio legato a forme di insegnamento specifico;
- 2) le categorie relative alle strutture della lingua in acquisizione si innestano, laddove possibile, su categorie preesistenti all'interno della competenza dell'apprendente. Gli assunti sviluppati in Krashen sono implicitamente corroborati da quanto postulato da Flege (v. § 2.2). In questo senso, possiamo concludere che analizzare la qualità delle categorie fonetiche già stabilitesi nella L1 sia fondamentale per prevedere gli effetti, a livello di rappresentazione mentale dei suoni, del contatto con una L2. Parallelamente, si ritiene necessario verificare (quantitativamente e qualitativamente) quanto la bontà dell'*input* di una L2 incida sull'acquisizione di suoni non-nativi in tale processo di contatto.

#### 2.4. Sistemi vocalici a confronto: italiano, romeno e tedesco

Come descritto da un'ampia letteratura sul tema (v. Loporcaro, Bertinetto, 2005; Krämer, 2009), l'inventario fonologico dell'italiano standard include sette fonemi vocalici in posizione tonica: /i, e, ε, a, ɔ, o, u/ (Figura 1). In posizione atona, invece, troviamo un sistema pentavocalico /i, e, a, o, u/, che si caratterizza per l'assenza delle due vocali medio-basse (anteriore e posteriore).

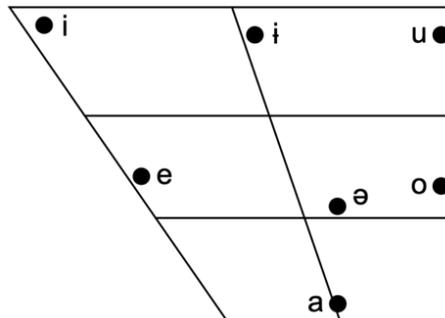
Figura 1. *Inventario vocalico dell'italiano* (Bertinetto, 2010: 7)



Attualmente, i principali riferimenti per la descrizione acustica dello spazio vocalico dell'italiano (standard/sovraregionale) sono gli studi di Ferrero (1968), Così, Ferrero, Vaggè (1995) e Loporcaro, Bertinetto (2005). Come sarà approfondito nel § 3.2, abbiamo basato l'analisi acustica delle vocali, sia dei locutori di italiano L2 sia del gruppo di controllo, sulla metodologia sperimentale impostata da Ferrero e colleghi per l'italiano, da Lambert Renwick (2012) per il romeno e Pätzold, Simpson (1997) per il tedesco.

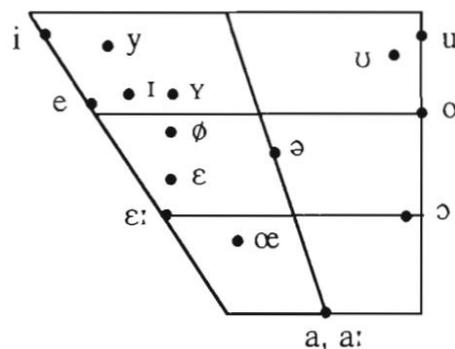
L'inventario del romeno comprende anch'esso sette monottonghi /i, i, u, e, ʌ, o, a/ (Figura 2) (Dészi, 2008: 438; Lambert Renwick, 2012: 132). Fonologicamente, il suo sistema vocalico è simmetrico: comprende tre vocali alte (anteriore, centrale e posteriore), tre vocali medie con gli stessi gradi di posteriorità e una vocale bassa centrale (Lambert Renwick, 2012: 148). Nonostante sia un sistema eptavocalico come l'italiano, il romeno si distingue da quest'ultimo per l'assenza del valore di opposizione per l'apertura vocale (/ɔ/ vs. /o/ e /ε/ vs. /e/) (Chitoran, 2002; Pană-Dindelegan, Maiden, 2013). Viceversa, sono presenti i suoni vocalici /ə/ e /i/, assenti in italiano standard.

Figura 2. *Inventario vocalico del romeno (Sarlin, 2014: 18)*



Il tedesco dispone di un inventario vocalico più ricco rispetto all'italiano, nonché di un tratto distintivo di lunghezza vocalica (es: *kam* [ka:m], pret. *kommen* 'venire' vs. *Kamm* [kam] 'pettine') (Figura 3). Come l'italiano e a differenza del romeno, il tedesco presenta l'opposizione tra medio-basse e medio-alte, sia sull'asse anteriore che sull'asse posteriore. Sia il tedesco che il romeno presentano la vocale centrale /ə/. Tuttavia, mentre il romeno e l'italiano presentano una corrispondenza quasi interamente trasparente tra grafia e fonia, il tedesco, pur essendo generalmente coerente dal punto di vista ortografico, presenta alcune eccezioni storicamente e normativamente motivate (es: *Blume* ['blu:mə] 'fiore', *Quelle* ['kvɛlə] 'sorgente' (Goswami, 2010).

Figura 3. *inventario vocalico del tedesco (Kohler, 1999: 87)*



In § 3.2 presenteremo i risultati preliminari dell'analisi acustica. Alla luce delle caratteristiche del sistema target L2 (italiano) in contatto con i sistemi L1 romeno/tedesco, forniremo quindi possibili interpretazioni riguardo l'efficacia nella produzione delle singole vocali L2, sulla base dello SLM.

### 3. ANALISI E RISULTATI

#### 3.1. *Descrizione del corpus ed elaborazione dati*

Le conclusioni ricomprese nel presente contributo derivano dall'analisi di una sottosezione del corpus CorAIIt (Corpus Audio di Italiano L2, Combei 2017). La scelta di tale sub-corpus (v. Tabella 1) è stata progettata in modo da garantire un equo

bilanciamento di tutte le principali variabili considerate nel corso della fase analitica (sesso, madrelingua, livello riferito di competenza per L2, modalità di acquisizione dell'L2, presenza di un training specifico per la pronuncia nel corso dell'acquisizione, *onset*, lunghezza della permanenza in Italia, lingua più usata in Italia).

Tabella 1. *Composizione del sub-corpus (Liv: livello riferito di competenza per la L2 nel quadro comune europeo; T.P.: training di pronuncia; Onset: precocità dell'onset, presto = 6-14 anni, tardi = dopo i 16 anni; Perm: permanenza in Italia, sempre = dalla nascita, lunga = 2-5 anni, media = 1-2 anni, breve = 6 mesi - 1 anno; L.I.: lingua più usata in Italia)*

	Sesso	Madrelingua	Liv.	Modalità	T.P.	Onset	Perm.	L.I.
1	M	Italiano	L1	spontanea	n.a.	nascita	sempre	Italiano
2	M	Italiano	L1	spontanea	n.a.	nascita	sempre	Italiano
3	M	Italiano	L1	spontanea	n.a.	nascita	sempre	Italiano
4	F	Italiano	L1	spontanea	n.a.	nascita	sempre	Italiano
5	F	Italiano	L1	spontanea	n.a.	nascita	sempre	Italiano
6	F	Italiano	L1	spontanea	n.a.	nascita	sempre	Italiano
7	M	Romeno	B1	spontanea	no	tardi	media	Romeno
8	M	Romeno	B2	spontanea	no	tardi	breve	Romeno
9	M	Romeno	C2	guidata + spontanea	sì	presto	lunga	Italiano
10	F	Romeno	C1	spontanea	no	tardi	media	Italiano
11	F	Romeno	C1	guidata + spontanea	sì	presto	breve	Romeno
12	F	Romeno	C2	guidata + spontanea	sì	presto	lunga	Italiano
13	M	Tedesco	B1	spontanea	no	tardi	breve	Tedesco
14	M	Tedesco	B2	guidata + spontanea	no	tardi	lunga	Tedesco
15	M	Tedesco	C2	guidata + spontanea	sì	presto	media	Italiano
16	F	Tedesco	B1	spontanea	no	tardi	breve	Tedesco
17	F	Tedesco	C1	guidata + spontanea	sì	presto	media	Italiano
18	F	Tedesco	B1	spontanea	no	tardi	breve	Italiano

Una delle variabili il cui ruolo testeremo è l'*onset*, termine con il quale si fa riferimento all'età della prima esposizione significativa del parlante all'italiano. È opportuno rimarcare che, da un lato, nel quadro teorico dell'Ipotesi del Periodo Critico (*Critical Period Hypothesis*, Lenneberg, 1967), l'età dell'*onset* è cruciale poiché è funzione delle possibilità acquisizioni dell'apprendente rispetto alla L2: in generale, si distinguono una fase di acquisizione precoce (*early bilingualism*), quando l'*onset* avviene nella fascia d'età compresa fra i 6 ed i 14 anni, e una fase di acquisizione tarda (*late bilingualism*) quando il primo contatto con la L2 si colloca dopo tale limite. Dall'altro, tale ipotesi è in aperto contrasto con quanto postulato da Flege, secondo il quale un'accurata produzione dei suoni non-nativi può essere raggiunta a qualsiasi età (quindi anche oltre la pubertà), in quanto dipendente dalla qualità della percezione dell'apprendente (Flege, 2005). Si è comunque ritenuto necessario raccogliere dati sull'età dell'*onset* al fine di verificare sperimentalmente se tale parametro fosse determinante al pari delle altre caratteristiche linguistiche e sociolinguistiche considerate – molte delle quali generalmente classificabili come indicatori della qualità e della quantità di *input* ricevuto (i.e. durata della permanenza in Italia, modalità di acquisizione dell'italiano, presenza/assenza di training di pronuncia, lingua più usata durante il soggiorno in Italia).

Per ciascuno dei diciotto soggetti, l'analisi ha riguardato un campione di parlato dalla durata orientativa di un minuto e venti (per un totale di 27'59" di parlato). Le singole campionature, ottenuti mediante la lettura di uno stesso, breve, articolo giornalistico, sono

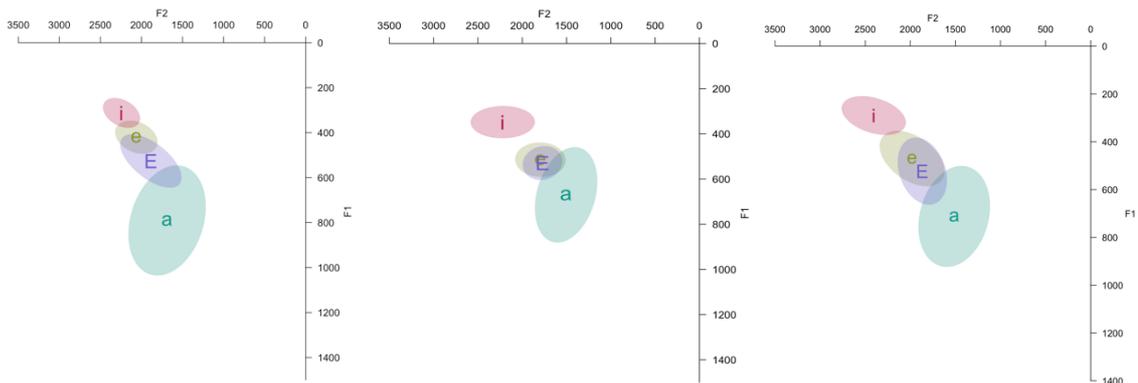
quindi state importate sul software Praat<sup>7</sup> (Boersma, Weenink, 2019) per la fase di segmentazione. L'operazione, avente come target le vocali toniche in posizione interconsonantica ed i suoni consonantici contestuali (consonante precedente / successiva), è stata completata attraverso l'annotazione manuale di tre *tiers*: uno per l'intera frase, uno per l'intera parola e uno per i singoli suoni target.

### 3.2. Analisi acustiche

Ai fini dell'analisi acustica sono stati estratti in maniera automatica i seguenti parametri: durata (in ms), frequenza fondamentale (f0), prima (F1), seconda (F2) e terza formante (F3) su 3 punti della traiettoria (25%, 50% e 75%)<sup>8</sup> per le sette vocali toniche in sillaba aperta all'interno di parole piane, con lo scopo di ridurre il più possibile la variabilità legata al contesto prosodico. Sulla scia dei lavori di Ferrero e colleghi, le ellissi di dispersione collocate su diagrammi cartesiani sono state repute il metodo più efficace per rappresentare lo spazio vocalico dei nostri locutori. Lo spazio è dato dalla correlazione tra la F1 (correlata al grado di apertura del tratto vocale nella produzione), la F2 (correlata al grado di altezza) e/o la F3 (correlata al grado di protrusione delle labbra) (Harrington, 2010). Le frequenze medie (in Hz) delle formanti di una data vocale rappresentano le coordinate del centro delle ellissi, mentre le dimensioni dei semi-assi delle ellissi sono proporzionali alla varianza congiunta stimata dei valori (ovvero, sono date da  $\pm 1$  errore standard) (Ferrero, 1968; Harrington, Cassidy, 1999)<sup>9</sup>. In base a questo tipo di rappresentazione, il valore medio di qualsiasi altro campione estratto dalla stessa popolazione ha quindi una data probabilità di cadere all'interno dell'ellissi.

Di seguito proponiamo alcuni grafici esemplificativi dei risultati dell'analisi acustica. In Figura 4 sono riportate ellissi di dispersione nello spazio F1xF2, calcolate per le vocali anteriori e suddivise per provenienza del campione maschile:

Figura 4. Ellissi di [i, e, ε, a], da sinistra rispettivamente: parlanti maschi italiani, romeni e tedeschi



Osserviamo che i valori della [a] sono visibilmente comparabili tra tutti i soggetti maschili per quanto riguarda la distanza relativa rispetto ai suoni adiacenti; ovvero, le sovrapposizioni che interessano la [a] sono marginali. Le [i] sono anch'esse molto simili,

<sup>7</sup> Il software è stato usato nella versione 6.0.56, disponibile alla pagina: <http://www.praat.org/>.

<sup>8</sup> Nel § 3.2 abbiamo considerato unicamente le formanti al 50% della durata complessiva della vocale, sostanzialmente corrispondente allo *steady state* del nucleo.

<sup>9</sup> L'area dell'ellissi è quindi misurata in Hz<sup>2</sup> (Calamai, 2003: 51).

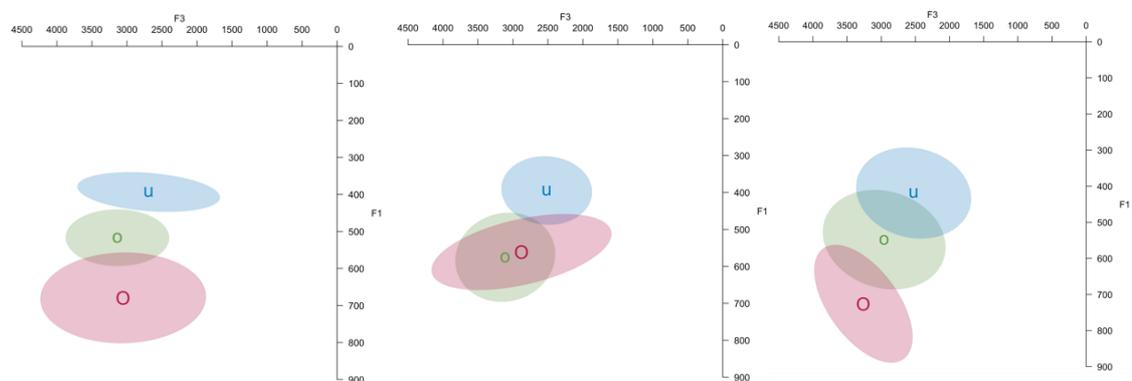
leggermente più alte nei tedescofoni, per cui la F2 assume valori fino a 2750 Hz. Notiamo invece che per i maschi italo-foni la [ɛ] (pur presentando parziali sovrapposizioni con le adiacenti [e] ed [a]) è ben definita, mentre si sovrappone in maniera più significativa alla vocale medio-alta anteriore nei maschi tedescofoni. La sovrapposizione è quasi totale nei romenofoni. Tale comportamento è del tutto in linea con quanto osservato da Chitoran (2002) e Pană-Dindelegan, Maiden (2013) sull'assenza del valore di opposizione fonologica per l'apertura vocalica /e - ɛ/ in romeno. Assumendo come riferimento lo SLM, possiamo infatti postulare che:

- i suoni in questione sono presumibilmente percepiti dai parlanti romenofoni di italiano L2 come non sufficientemente dissimili l'uno dall'altro per attivare la creazione di una nuova categoria fonetica;
- si attiva quindi il meccanismo di “*equivalence classification*”, per il quale le categorie in questione subiscono un “*merge*” a livello di proprietà acustiche (§ 2.2);

L'effetto di *merge* è ben evidente nei valori formantici. Per i romeni, la F1 e la F2 di entrambe le vocali medie anteriori sono rispettivamente comprese tra i 450 e 550 Hz e 1500 e 2000 Hz. Di contro, i valori formantici di F1 e F2 negli italiani sono visibilmente distinti: abbiamo rispettivamente 350-450 Hz e 1750-2400 Hz per la [ɛ], e 400-600 Hz e 350-500 Hz per la [e]. Per quanto riguarda il gruppo tedescofono, come già accennato, la sovrapposizione delle due categorie non è completa. Tuttavia, a livello impressionistico, le [e] sono state percepite da tutti gli autori come più simile alla [ə] in posizione finale di parola. L'effetto potrebbe essere dovuto, nel caso dei parlanti tedescofoni meno esperti, alla mancanza di corrispondenza biunivoca tra grafema <e> e fonema /e/ in posizione finale, come già suggerito da Kohler (1999). Avendo analizzato nel presente lavoro solo le vocali toniche, questi valori non sono riportati nei grafici. Si propone quindi di condurre in futuro studi sulle vocali atone in posizione finale, al fine di confermare o meno l'effetto di centralizzazione anche a livello acustico. Per quanto riguarda la collocazione delle vocali anteriori, si è osservata una tendenza alla sovrapposizione di medio-basse e medio-alte anche nel sottogruppo femminile; il *pattern* è del tutto comparabile a quello del sottogruppo maschile, anche se meno marcato.

Le seguenti ellissi riportano lo spazio F1xF3 occupato dalle vocali posteriori, suddivise per provenienza del campione femminile:

Figura 5. Ellissi di [ɔ, o, u], da sinistra rispettivamente: parlanti femmine italiane, romene e tedesche



Si osserva dai grafici che le vocali posteriori sono nettamente distinte per grado di protrusione (F3) nelle italofone, in maniera assolutamente evidente nelle romenofone e

mediamente nelle tedescofone. Lo stesso avviene nei parlanti maschi romeni e tedeschi: complessivamente, notiamo che il campione maschile esibisce più variazione nella produzione a livello intra-gruppo. Per quanto riguarda l'evidente coincidenza /o - ɔ/ nell'intero gruppo romenofono, abbiamo quindi uno scenario simile a quello già osservato per /e - ε/. Di nuovo, assistiamo ad un possibile effetto di *category assimilation*; anche la vocale posteriore medio-bassa non è infatti distintiva nell'inventario vocalico romeno.

### 3.3. Interpretazione statistica

Partendo dai risultati delle analisi acustiche e usando un modello di regressione lineare per ogni singola vocale, abbiamo testato l'effetto delle variabili sociolinguistiche dei locutori (i.e. il genere, la L1, il livello di competenze linguistiche in italiano, il tipo di apprendimento linguistico, onset-l'età della prima esposizione all'italiano, la presenza o l'assenza delle lezioni di pronuncia per l'italiano, il periodo di tempo trascorso in Italia e la lingua usata durante questa permanenza) sulla qualità della produzione delle vocali italiane da parte dei soggetti non-nativi. La variabile dipendente è F2-F1 per le vocali anteriori e F3-F1 per le vocali posteriori, indicatori usati frequentemente per questo scopo nella fonetica sperimentale (v. Fant, 1983). Sebbene verranno presentati i risultati ottenuti per tutte le vocali, si è scelto di analizzare in dettaglio solo quelli che mostrano *pattern* statisticamente significativi.

Per quanto riguarda le vocali [a], [e], [o] e [u] i valori medi sono simili per tutte le tipologie di parlanti ([a] ha un rapporto F2-F1 medio di circa 850 Hz, per [e] il valore medio è vicino a 1700 Hz, per [o] il valore medio di F3-F1 si attesta attorno a 3000 Hz, mentre per [u] attorno a 2100 Hz). La sottile differenza nelle produzioni di alcuni parlanti, in particolar modo per i tedeschi e generalmente per la [e] e per la [u], non è statisticamente significativa; dal momento che questo comportamento non può essere giustificato da nessuna delle variabili sociolinguistiche considerate, riteniamo che si tratti piuttosto di una variazione idiosincratica.

Nel caso della [i] emergono alcune differenze statisticamente significative. L'*output* del modello di regressione lineare può essere riassunto come segue: osservazioni = 365,  $F(11,353) = 41,50$ ,  $p = 0,00$   $R^2 = 0,56$  Adj.  $R^2 = 0,55$ . La L1 dei parlanti non spiega la differenza tra i valori registrati per questa vocale ( $p\text{-value} > 0,05$ ), mentre l'effetto è statisticamente significativo per la durata della permanenza in Italia ( $p\text{-value} < 0,01$ ) e per la lingua predominante durante il soggiorno ( $p\text{-value} < 0,01$ ). Generalmente i parlanti con una lunga permanenza in Italia e che utilizzano quotidianamente l'italiano sono in grado di produrre una [i] più simile a quella dei parlanti madrelingua.

Anche per quanto riguarda le vocali semi-aperte emergono alcuni *pattern* interessanti. Riassumiamo di seguito i risultati del modello per la vocale [ɛ]: osservazioni = 186,  $F(8,177) = 6,71$ ,  $p = 0,00$   $R^2 = 0,23$  Adj.  $R^2 = 0,20$ . Le variabili che possono predire differenze significative nella produzione della vocale anteriore semi-aperta sono la L1 dei parlanti ( $p\text{-value} = 0,01$ ), il livello di competenza linguistica in italiano ( $p\text{-value} < 0,01$ ), la presenza del training di pronuncia ( $p\text{-value} = 0,04$ ) e la lingua predominante durante la permanenza in Italia ( $p\text{-value} < 0,01$ ). Come mostrato graficamente in § 3.2 tramite le ellissi, la realizzazione di [ɛ] da parte di diversi apprendenti di italiano L2, specialmente romeni, si distanzia notevolmente da quella dei parlanti italiani. Il più delle volte i valori della [ɛ] coincidono con quelli della [e]. Da un'attenta analisi dei coefficienti del modello statistico sembra che gli unici soggetti in grado di produrre accuratamente la [ɛ] siano i parlanti romeni che hanno un ottimo livello di italiano (i.e. C2), che hanno ricevuto lezioni di pronuncia, e che usano principalmente l'italiano per comunicare.

Infine, l'adattamento del modello di regressione lineare per la F3-F1 della [ɔ] è il seguente: osservazioni = 36,  $F(7,28) = 30,94$ ,  $p = 0,00$   $R^2 = 0,89$  Adj.  $R^2 = 0,86$ . Le variabili che mostrano un effetto significativo sulla bontà della produzione della vocale posteriore semi-aperta sono la L1 dell'apprendente ( $p\text{-value} < 0,01$ ), la durata della permanenza in Italia ( $p\text{-value} = 0,02$ ) e la principale lingua usata per comunicare ( $p\text{-value} = 0,04$ ). Generalmente, i parlanti tedescofoni producono una [ɔ] del tutto comparabile a quella degli italiani, mentre per i parlanti romeni emerge una sovrapposizione quasi perfetta con la [o]. Tuttavia, gli apprendenti romenofoni che usano spesso l'italiano e che vivono in Italia da più di 48 mesi riescono a discriminare l'opposizione /o - ɔ/, anche se essa non ha valore fonemico nella loro L1.

### 3.4. Clustering

In questo studio proponiamo anche l'utilizzo di *k-Means*, un algoritmo euristico di *clustering* partizionale ispirato ai lavori di Forgy (1965) e MacQueen (1967). Il funzionamento del *k-Means* può essere riassunto come segue: nell'*input* si riceve un parametro *k*; l'insieme di *n* oggetti da analizzare viene ripartito in *k* cluster in modo da ottenere un'alta similarità intra-*cluster* e una bassa similarità inter-*cluster*; questa similarità si calcola rispetto al valore medio degli oggetti contenuti nel *cluster*.

Per gli esperimenti di *clustering* degli apprendenti abbiamo usato le funzioni disponibili sulla libreria *Orange*<sup>10</sup> di *Python* (Demsar *et al.*, 2013). Considerando che le vocali /a/, /e/, /o/ e /u/ vengono prodotte da nativi e non-nativi in una maniera tutto sommato simile, e per esigenze di spazio abbiamo ritenuto opportuno limitarci a esaminare soltanto le vocali semi-aperte. Per caratterizzare le [ɛ] e le [ɔ] di ognuno dei 18 locutori abbiamo usato i seguenti 11 parametri: la durata della vocale, la f0, i valori di F1, F2 e F3 nei tre punti descritti in § 3.2. Con lo scopo di ottenere il numero ideale di *cluster* in base alle caratteristiche delle vocali dei parlanti, abbiamo usato la funzione *silhouette* che genera punteggi da -1 a 1 per ogni potenziale *k*. In seguito, il *k* è stato usato come *input* nell'algoritmo di *clustering* per raggruppare i parlanti per similarità. Per visualizzare l'*output* del *clustering* abbiamo usato un grafico di tipo *Multidimensional Scaling* (MDS) ottenuto tramite il *Principal Component Analysis* (Jolliffe, 2007) dello spazio vettoriale dei parametri. Questa visualizzazione ci permette anche di inserire *metadati*, nel nostro caso le variabili sociolinguistiche dei parlanti.

Per quanto riguarda la vocale [ɛ], il numero ideale di cluster è 2, con un punteggio di 0.72, suggerendo una distinzione quasi netta tra due tipologie di produzioni. Infatti, dal grafico MDS (v. Figura 6) si evince chiaramente che a destra troviamo solo parlanti romeni, mentre a sinistra troviamo tutti i parlanti italiani e tedeschi e due parlanti romeni. Le caratteristiche sociolinguistiche dei due apprendenti romeni che condividono il *cluster* con gli italiani e i tedeschi indicano un'altissima competenza linguistica in italiano e una permanenza lunga in Italia.

Infine, anche nel caso della [ɔ], il numero ideale di cluster è 2, con un punteggio di 0.62. Come evidenziato nel grafico MDS (v. Figura 7), la pronuncia della vocale semi-aperta arrotondata nella maggior parte dei parlanti romeni denota un'ampia dispersione e un distanziamento visibile da quella dei parlanti italiani e della maggior parte dei parlanti tedeschi. Inoltre, si può constatare che gli apprendenti romeni e tedeschi che hanno un livello alto di competenze linguistiche in italiano, che hanno ricevuto lezioni di pronuncia

<sup>10</sup> La documentazione relativa alla libreria Orange è disponibile online a questa pagina: <https://docs.biolab.si//3/data-mining-library/>.

e che hanno soggiornato in Italia per lunghi periodi producono una [ɔ] del tutto comparabile a quella dei locutori italiani; questo dato è rilevante per l'insegnamento dell'italiano ad apprendenti stranieri.

Figura 6. *MDS dei valori medi di [ɛ] di ogni parlante*

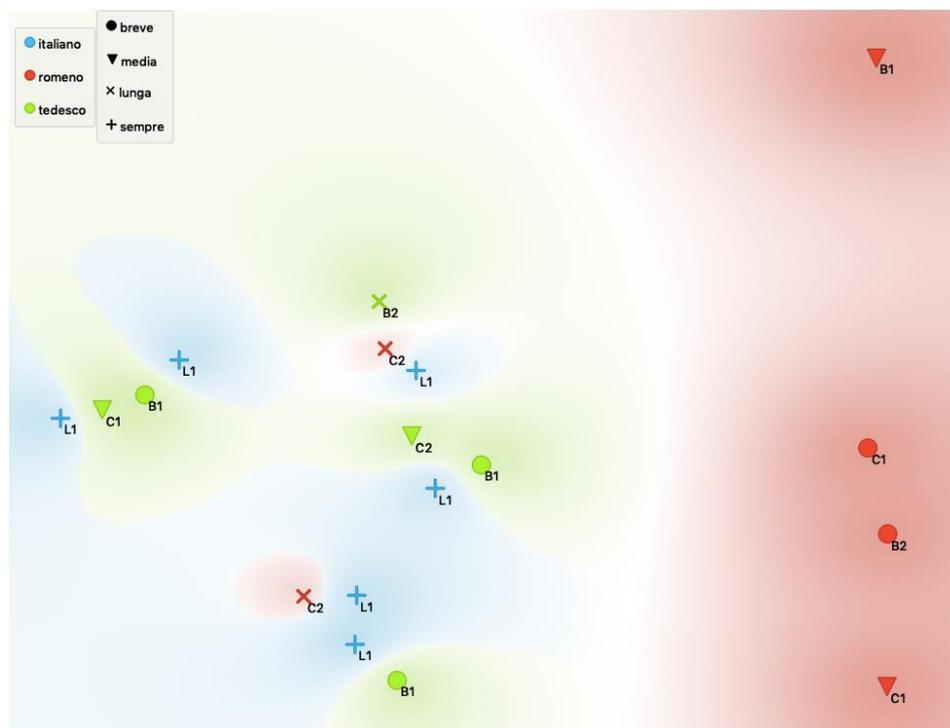
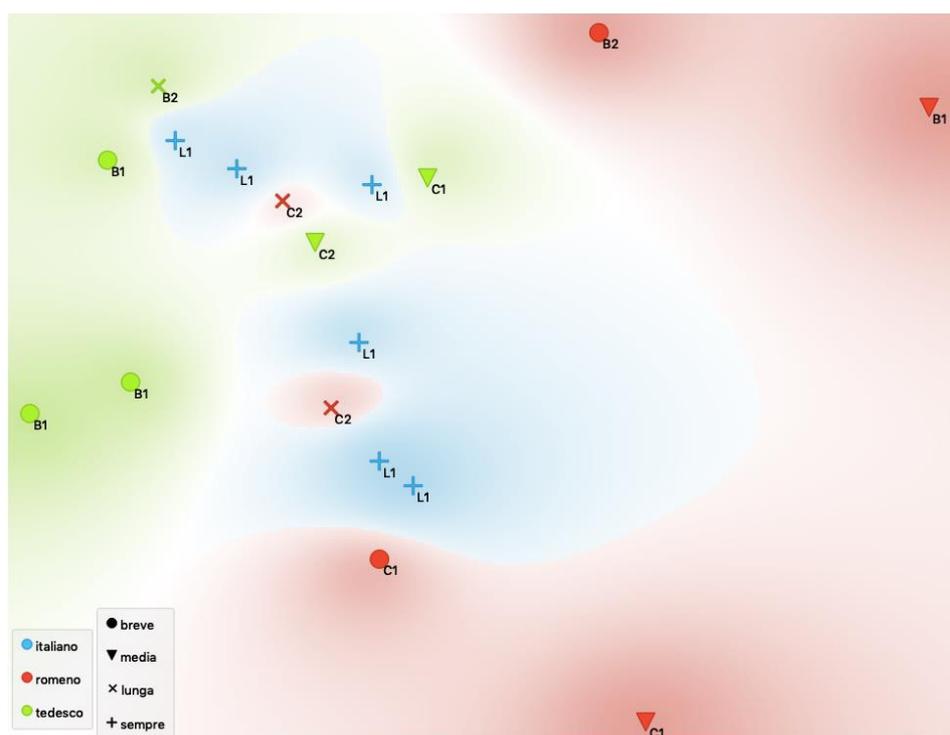


Figura 7. *MDS dei valori medi di [ɔ] di ogni parlante*



#### 4. CONCLUSIONI

Questo lavoro sperimentale ha avuto come oggetto di indagine il vocalismo tonico di 12 apprendenti adulti di italiano L2 provenienti dalla Germania e dalla Romania che avevano soggiornato in Italia per un periodo pari o superiore ai 6 mesi al momento della registrazione. I risultati dimostrano l'effetto rilevante dell'*equivalence classification* (Flege, 1995, 2007) nel caso della pronuncia delle vocali semi-aperte italiane da parte degli apprendenti romenofoni. Analizzando queste produzioni, appare evidente come il grado di apertura risulti spesso semi-chiuso. Questo ci indica che alcuni locutori romeni non riescono a discriminare le opposizioni /e - ε/ e /o - ɔ/, d'altronde contrasti senza valore fonemico nella loro L1. Gli apprendenti tedeschi, invece, non riscontrano difficoltà nella produzione delle vocali semi-aperte. A livello acustico, per alcuni locutori tedeschi è stata evidenziata una pronuncia atipica delle /e/ e /u/ italiane, anche se queste differenze non sono statisticamente significative. Attraverso modelli statistici e tecniche di *clustering* si è potuto osservare che generalmente una pronuncia più simile a quella dei parlanti madrelingua si riscontra nel caso degli apprendenti che vivono da più di 48 mesi in Italia, che usano principalmente l'italiano per comunicare, che hanno ricevuto lezioni esplicite di pronuncia e che hanno un livello alto di competenze linguistiche in italiano. L'età della prima esposizione all'italiano non risulta statisticamente significativa. Questo dato è rilevante per la didattica dell'italiano L2 poiché indica una rilevanza maggiore della qualità e della quantità dell'*input* nell'apprendimento dei tratti segmentali, come suggerito da altri studiosi (Flege, 2009; Camus, 2019; v. § 3.1). Le future direzioni di lavoro riguarderanno l'analisi strumentale delle vocali atone, del consonantismo e dei tratti prosodici, allo scopo di indagare al meglio l'acquisizione della fonologia dell'italiano L2.

#### RIFERIMENTI BIBLIOGRAFICI

- Bergmann C., Nota A., Sprenger S.A., Schmid M.S. (2016), "L2 immersion causes non-native-like L1 pronunciation in German attriters", in *Journal of Phonetics*, 58, pp. 71-86.
- Bertinetto P. M. (2010), "Fonetica italiana", in *Quaderni del Laboratorio di Linguistica della Scuola Normale Superiore*, 9, 1, pp. 1-30.
- Calamai S. (2003), "Vocali fiorentine e vocali pisane a confronto", in Albano Leoni F., Pettorino M., Savy R. (a cura di), *Proceedings of the National Conference: Il parlato Italiano*, D'Auria, Napoli, pp. 1-25.
- Camus P. (2019), "The effects of explicit pronunciation instruction on the production of second language Spanish voiceless stops: a classroom study", in *Instructed Second Language Acquisition*, 3, 1, pp. 81-103.
- Chang B. C. (2012), "Rapid and multifaceted effects of second-language learning on first-language speech production", in *Journal of Phonetics*, 40, pp. 249-268.
- Chițoran I. (2002), *The Phonology of Romanian: A Constraint-Based Approach*, Mouton de Gruyter, Berlino-New York.
- Combei C. R., Marotta G. (2019), "Where do accents come from? Factors affecting the degree of foreign-accented Italian", in Piccardi D., Ardolino F., Calamai S. (a cura di), *Gli archivi sonori al crocevia tra scienze fonetiche, informatica umanistica e patrimonio digitale*, Officinaventuno, Milano, pp. 233-248:  
[http://www.aisv.it/StudiAISV/2019/vol\\_6/014\\_CombeiMarotta.pdf](http://www.aisv.it/StudiAISV/2019/vol_6/014_CombeiMarotta.pdf).

- Combei C. R. (2017), “CorAIIt – A non-native speech database for Italian”, in Basili R., Nissim M., Satta G. (a cura di), *Proceedings of the Fourth Italian Conference on Computational Linguistics*, Accademia University Press, Torino, pp. 113-118.
- Combei C. R. (2019), *A corpus-based study on foreign-accented Italian*, tesi di dottorato, Università di Pisa, Pisa.
- Così P., Ferrero F., Vagges K. (1995), “Rappresentazioni acustiche e uditive delle vocali italiane”, in Cocchi A. (a cura di), *Proceedings of the 23<sup>rd</sup> ALA National Conference*, 12-14 August 1995, Bologna, pp. 151-156.
- Costamagna L. (2008), “Acquisition of Italian L2 affricates: the case of a Brazilian learner”, in Rauber A. S., Watkins M. A., Baptista B. O. (a cura di), *New Sounds 2007: Proceedings of the Fifth International Symposium on the Acquisition of Second Language Speech*, Federal University of Santa Catarina, Florianópolis, pp. 138-148.
- de Leeuw E. (2008), “When your native language sounds foreign: a phonetic investigation into first language attrition”, tesi di dottorato, Queen Margaret University, Edinburgh.
- De Meo A, Vital M., Xu Y. (2015), “Lo sviluppo del sistema vocalico dell’italiano nell’interlingua di adulti sinofoni”, in Vayra M., Avesani C., Tamburini F. (a cura di.), *Il farsi e il disfarsi del linguaggio*, Officinaventuno, Milano, pp. 191-202:  
[http://www.aisv.it/StudiAISV/2015/vol\\_1/012\\_DeMeoVitalXu.pdf](http://www.aisv.it/StudiAISV/2015/vol_1/012_DeMeoVitalXu.pdf).
- Demsar J., Curk T., Erjavec A., Gorup C., Hocevar T., Milutinovic M., Mozina M., Polajnar M., Toplak M., Staric A., Stajdohar M., Umek L., Zagar L., Zbontar J., Zitnik M., Zupan B. (2013), “Orange: Data Mining Toolbox in Python”, in *Journal of Machine Learning Research*, 14, pp. 2349-2353.
- Dészi, M. (2008), “The morphology of Romanian palatalization”, in *Verbum Analecta Neolatina*, 10, 2, pp. 437-448.
- Fant G. (1983), “Feature analysis of Swedish vowels - a revisit”, in *STL-QPSR KTH*, 2-3, pp. 1-19.
- Ferrero F. (1968), “Diagrammi di esistenza delle vocali italiane”, in *Alta frequenza*, 37, pp. 54-58.
- Ferrero F., Magno Caldognetto E., Così P. (1996), “Sui piani formantici acustici e uditivi delle vocali di uomo, donna, bambino”, in Peretti A., Simonetti P. (a cura di), *Proceedings of the 24th ALA National Conference*, Arti Grafiche Padovane, Padova, pp. 169-178.
- Flege J. E. (1995), “Second language speech learning: Theory, findings, and problems”, in Strange W. (a cura di), *Speech perception and linguistic experience: Issues in Cross-Language Research*, York Press, Baltimore, pp. 233-277.
- Flege J. E. (2007), “Language contact in bilingualism: Phonetic system interactions”, in Cole J., Hualde J. (a cura di), *Laboratory Phonology 9*, Mouton de Gruyter, Berlino-New York, pp. 353-380.
- Flege J. E. (2009), “Give input a chance!” in Piske T., Young-Scholten M. (a cura di), *Input matters in SLA*, Multilingual Matters, Bristol, pp. 175-190.
- Forgy E. W. (1965), “Cluster analysis of multivariate data: efficiency versus interpretability of classifications”, in *Biometrics*, 21, 3, pp. 768-769.
- Goswami U. (2010), “Phonology, reading, and reading difficulties”, in Hall K., Goswami U., Harrison C., Ellis S., Soler J. (a cura di), *Interdisciplinary Perspectives on Learning to Read*. Routledge, London-New York pp. 103-116.
- Harrington J. (2010), *The phonetic analysis of speech corpora*, Wiley-Blackwell, Chichester.
- Harrington J., Cassidy S. (1999), *Techniques in speech acoustics*, Springer Verlag, Dordrecht.
- Jolliffe I. T. (2007), *Principal component analysis*, Springer, New York.

- Kohler K. (1999), "German", in *Handbook of the Phonetic Association*, Cambridge University Press, Cambridge, pp. 86-89.
- Krämer M. (2009), *The Phonology of Italian*, Oxford University Press, Oxford.
- Krashen S. D. (1981), *Second Language Acquisition and Second Language Learning*, Pergamon, Oxford.
- Krashen S. D. (1985), *The Input Hypothesis: Issues and Implications*, Longman, New York.
- Lambert Renwick M. E. (2012), *Vowels of Romanian: Historical, Phonological and Phonetic Studies*, Tesi di dottorato, Cornell University, Ithaca, NY.
- Lenneberg E. H. (1967), *The Biological Foundations of Language*, Wiley and sons, New York.
- Lin C. (2013), "Perception and Production of English Front Vowels by Taiwanese EFL Learners", in *Theory and Practice in Language Studies*, 3, 11, pp. 1952-1958.
- Loporcaro M., Bertinetto P.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, pp. 131-151.
- MacQueen J. (1967), "Some methods for classification and analysis of multivariate observations", in *Proceedings of the Fifth Berkeley Sympos. Math. Statist. and Probability*, I: Statistics, University California Press, Berkeley, pp. 281-297.
- Major R. C. (2001), *Foreign accent: The ontogeny and phylogeny of second language phonology*, Lawrence Erlbaum Associates Publishers, Mahwah.
- Munro M. (1998), "The effects of noise on the intelligibility of foreign-accented speech", in *Studies in Second Language Acquisition*, 20, pp. 139-154.
- Pană-Dindelegan G., Maiden M. (2013), *The grammar of Romanian*, Oxford University Press, Oxford.
- Pätzold M., Simpson A. P. (1997), "Acoustic analysis of German vowels in the Kiel Corpus of Read Speech", in Köhler K. J. (a cura di), *Arbeitsberichte des Instituts für Phonetik und Digitale Sprachverarbeitung der Universität Kiel*, Universität Kiel, Kiel, pp. 215-247.
- Pellegrino E. (2012), "The perception of foreign accented speech. Segmental and suprasegmental features affecting degree of foreign accent in Italian L2", in Mello H., Pettorino M., Raso T. (a cura di), *Proceedings of the VIIth GSCP International Conference - Speech and Corpora*, Firenze University Press, Firenze, pp. 261-267.
- Piske T., Flege J. E., MacKay I., Meador D. (2011), "Investigating native and non-native vowels produced in conversational speech", in Wrembel M., Kul M., Dziubalska-Kolaczyk K. (a cura di), *Achievements and perspectives in the acquisition of second language speech: New Sounds 2010*, 2, Peter Lang, Berna, pp. 195-205.
- Romito L., Frontera M. (2015), "Perception and production of Italian L2 sounds", in Botinis A. (a cura di), *Proceedings of the International Conference of Experimental Linguistics ExLing 2015*, National and Kapodistrian University of Athens, Atene, pp. 70-73.
- Sarlin M. (2014), *Romanian Grammar*, Books on Demand, Helsinki.
- Sereno J. (2016), "The relative contribution of segments and intonation to the perception of foreign-accented speech", in *Applied Psycholinguistics*, 2, 37, pp. 303-322.
- Vayra M., Avesani C., Best C., Bohn O. (2012), "Non solo dettagli fonetici, non solo categorie fonologiche: l'interazione tra fonetica e fonologia nella percezione di suoni non-nativi", in *Studi e saggi linguistici*, 50, 2, pp. 119-145.