

# ACQUISIZIONE DELLE VOCALI L2 DELL'INGLESE AMERICANO DA PARTE DI APPRENDENTI ITALOFONI: UN CONFRONTO TRA TRAINING PERCETTIVO E TRAINING ARTICOLATORIO CON ECOGRAFO

IMMACOLATA SONIA D'APOLITO, BIANCA SISINNI, MIRKO GRIMALDI,  
BARBARA GILI FIVELA  
UNIVERSITÀ DEL SALENTO

**Abstract** – La realizzazione del contrasto vocalico dell'inglese americano /ɑ-ʌ/ (es. *cop-cup*) da parte di apprendenti italofoeni risulta essere particolarmente difficile in produzione, poiché le differenze fonetiche-fonologiche dei due sistemi linguistici e le regole di conversione grafema-fonema fanno sì che gli apprendenti italofoeni realizzino il contrasto L2-/ɑ-ʌ/ utilizzando le vocali native /ɔ-a/ rispettivamente. Lo scopo di questo studio è quello di osservare se un breve *training* possa aiutare gli apprendenti italofoeni a migliorare la produzione delle vocali non native /ɑ-ʌ/. In particolare, si effettuerà un confronto tra un *training* percettivo e un *training* articolatorio, effettuato con l'utilizzo di un ecografo, al fine di osservare: 1) gli effetti di entrambi i *training* sulla realizzazione del contrasto L2 /ɑ-ʌ/; e 2) se il *training* articolatorio comporti risultati migliori rispetto a quello percettivo. Allo studio hanno partecipato nove studentesse salentine che sono state suddivise, in modo casuale, in tre gruppi: i) tre soggetti hanno preso parte al *training* percettivo della durata di un'ora (ES-P); ii) tre soggetti hanno partecipato al *training* articolatorio con ultrasuono della durata di un'ora (ES-US); iii) tre soggetti di controllo che non hanno effettuato alcun tipo di *training* (CS). I primi due gruppi (ES-P e ES-US) hanno ricevuto istruzioni sulle caratteristiche fonetiche sia delle vocali L2 /ɑ-ʌ/ che delle vocali della L1 /ɔ-a/ nonché sulle loro differenze utilizzando una rappresentazione sul piano cartesiano delle formanti F1 e F2. Successivamente, ciascun gruppo ha effettuato il *training*: i) i soggetti del gruppo ES-P hanno effettuato un test di identificazione basato sulla procedura *High Variability Phonetic Training*; ii) i soggetti del gruppo ES-US hanno effettuato un *training* articolatorio osservando sia un video relativo ai gesti linguistici durante la produzione delle vocali /ɑ-ʌ/ da parte di un parlante nativo, sia una visualizzazione dinamica dei movimenti e della posizione della lingua, ottenuta grazie ad una sonda ecografica posizionata sotto il proprio mento. Le partecipanti sono state registrate sia prima che dopo l'addestramento (pre- e post-test) e le loro produzioni sono state analizzate acusticamente misurando il valore delle prime due formanti. I risultati mostrano che nel pre-test tutti le apprendenti realizzano L2-/ɑ-ʌ/ come L1-/ɔ-a/ rispettivamente. Al contrario, nel post-test sono soprattutto le apprendenti che hanno effettuato il *training* articolatorio riescono a realizzare le vocali L2 in modo differente sia rispetto a quelle realizzate nel pre-test sia rispetto alle vocali della L1. In particolare, un'apprendente produce entrambe le vocali non native in modo più accurato. Il *training* articolatorio sembra quindi essere più efficace del *training* percettivo.

**Keywords:** produzione vocali L2, *training* percettivo, *training* articolatorio, ecografo, ultrasuoni.

## 1. Introduzione

Nonostante la pronuncia sia riconosciuta come una competenza importante nell'apprendimento di una seconda lingua, il suo insegnamento in un contesto formale risulta, invece, ampiamente sottovalutato all'interno dei programmi scolastici (Kirkova-Naskova, 2019). Il ruolo secondario adibito alla pronuncia deriva, da un lato, dalla scelta del docente di non insegnare la pronuncia al pari delle altre abilità linguistiche e,

dall'altro, dalla mancanza di materiale didattico specifico che aiuti a migliorare la pronuncia. Gli insegnanti nativi molto spesso non hanno un'adeguata preparazione all'insegnamento della pronuncia della loro lingua materna e, oltretutto, nel tempo dimostrano di avere un atteggiamento di tolleranza verso gli errori di pronuncia; gli insegnanti che, invece, non sono parlanti nativi della lingua che insegnano sono molto spesso insicuri della propria pronuncia e prestano, quindi, maggiore attenzione ad insegnare e a far esercitare gli studenti su altre competenze linguistiche (Henderson *et al.* 2012; Kirkova-Naskova *et al.* 2013). Ad esempio, studi condotti da Brown (1992), Fraser (2000) e Yates (2001) hanno dimostrato, infatti, che molti insegnanti tendono ad evitare di insegnare in modo esplicito la pronuncia per mancanza di sicurezza, abilità e conoscenza. Inoltre, questi studi dimostrano che il metodo di insegnamento e la mancanza di materiali appropriati provocano inadeguatezze tanto nell'insegnamento della pronuncia quanto nell'apprendimento (Hismanoglu, Hismanoglu 2010). Tutto ciò non fornisce agli studenti le conoscenze necessarie a livello fonetico-fonologico, né a livello segmentale né prosodico e, di conseguenza, gli apprendenti non riescono a superare gli errori di pronuncia, a migliorare l'accuratezza, la fluenza, l'ascolto e a raggiungere la piena autonomia e una maggiore sicurezza nel comunicare in una lingua L2 (Hismanoglu, Hismanoglu 2010). Ai fini dell'insegnamento della pronuncia sono stati individuati, secondo Celce-Murcia, Brinton e Goodwin (2007), tre approcci: a) approccio intuitivo-imitativo; b) approccio analitico-linguistico; e c) un approccio integrativo. In base all'approccio intuitivo-imitativo, gli apprendenti possono acquisire in modo implicito il sistema fonologico della lingua L2 sulla base dell'ascolto e dell'imitazione della prosodia e dei suoni della L2, senza ricevere alcuna informazione esplicita sulla pronuncia. Gli apprendenti devono essere esposti ad un buon modello anche attraverso il supporto di video, siti web, software, ecc. In base all'approccio analitico-linguistico, gli apprendenti ricevono indicazioni specifiche sulla pronuncia attraverso grafici con i simboli fonetici, informazioni sull'apparato di fonazione e descrizioni dei movimenti articolatori, informazioni contrastive tra la L1 e L2 ecc. Anche in questo caso si possono utilizzare strumenti interattivi come software e siti web (Lee 2008). Le informazioni esplicite permettono all'apprendente di focalizzare l'attenzione sui suoni e sulla prosodia della lingua L2. Questo approccio nasce per affiancare e completare l'approccio intuitivo-imitativo. Infine, l'approccio integrativo è legato alla comunicazione e prevede una serie di attività di ascolto orientate alla pronuncia in modo da facilitarne l'apprendimento. Il focus e il beneficio sono duplici in quanto riguardano un micro-livello, che si basa sulla competenza linguistica (fonetica e fonologica) attraverso l'esercitazione sul piano segmentale e soprasegmentale, ed un macro-livello che mira alla comunicazione, con lo scopo di sviluppare le competenze e le strategie attraverso l'uso della lingua per scopi comunicativi (Elliot 1997). In sostanza, la pronuncia viene insegnata per andare incontro alle necessità dell'apprendente.

Un metodo di insegnamento non focalizzato sulla pronuncia, come anche uno scarso uso della lingua L2, comporta un impoverimento della corretta percezione dei suoni L2 e di conseguenza della loro produzione, essendo la percezione un fattore importante al fine della categorizzazione e discriminazione dei fonemi L2 (Best, Tyler 2007). Nonostante ciò, i risultati degli studi sulla percezione dei suoni L2 da parte di adulti mostrano che i percipienti non nativi possono percepire senza difficoltà alcuni contrasti della L2 e, al contrario, mostrare difficoltà nel percepirne altri. Ciò dipende dalle caratteristiche fonetiche e fonologiche della lingua madre (L1), ossia in base a quanto i suoni della L2 sono simili o dissimili rispetto ai suoni della L1 (Best, Tyler 2007). La pronuncia di una seconda lingua risulta essere, quindi, ampiamente influenzata sia dalla

lingua materna che dalle caratteristiche della L2, perché entrambe influenzano la percezione. Gli apprendenti di una seconda lingua possono però migliorare la propria pronuncia dei suoni non nativi attraverso un training, un addestramento mirato. I *training* possono essere di tipo percettivo e/o articolatorio o una combinazione dei due. I *training* percettivi hanno l'obiettivo di migliorare la percezione dei contrasti L2 difficili, in particolare la percezione di quei suoni che vengono assimilati all'interno di una categoria della L1 (Best 1995). Il compito è quello di addestrare e migliorare la capacità di percezione attraverso test di discriminazione e/o di identificazione. Nel test di identificazione, i soggetti visualizzano sullo schermo di un computer una coppia minima (es. *rock-lock*) ed ascoltano un solo stimolo. Dopo l'ascolto, devono cliccare sulla parola che a loro sembra corrispondere allo stimolo audio e ricevono un messaggio di *feedback*, ossia "corretto" se la risposta è corretta e "sbagliato" per la risposta sbagliata (di solito con la possibilità di riascoltare lo stimolo stesso). Nel test di discriminazione, i soggetti devono stabilire se i suoni proposti siano uguali o differenti. Nel test AX, ad esempio, ai soggetti vengono presentati due stimoli A e X e il loro compito è quello di dire se X sia uguale ad A oppure no. Molti studi sono relativi a *training* di tipo percettivo il cui obiettivo è osservare non solo un miglioramento della percezione dei suoni L2, ma anche l'efficacia rispetto alla produzione dei suoni L2. Studi sperimentali hanno infatti dimostrato che la pronuncia dei suoni L2 migliora anche grazie a *training* di tipo percettivo, in particolar modo con il metodo *High Variability Phonetic Training* (HVPT) (Logan *et al.* 1991; Bradlow *et al.* 1997; Iverson, Evans 2009; Ylinen *et al.* 2010). Questo metodo si basa sull'utilizzo di un test di identificazione nel quale l'attenzione degli apprendenti è focalizzata su determinate proprietà fonetiche rilevanti per categorizzare i suoni L2. Il suono bersaglio viene presentato in diversi contesti fonetici (variabilità degli stimoli) derivanti da produzioni naturali da parte di diversi parlanti (variabilità del parlante). Questa alta variabilità permette ai soggetti la formazione di una nuova e robusta categoria fonetica (Lively *et al.* 1994) e la generalizzazione a nuovi stimoli e a nuovi parlanti (Posner, Keele, 1968, Barriuso, Hayes-Harb 2018). Il metodo HVPT ha un effetto positivo anche sulla produzione dei suoni L2, come dimostrato in Bradlow *et al.* (1997), in quanto l'apprendimento percettivo è strettamente collegato all'apprendimento in produzione sebbene i due processi potrebbero anche essere distinti per ciascun soggetto (Bradlow *et al.* 1997). Infatti, uno dei più importanti fattori da considerare è la variazione tra i soggetti nelle loro performance sia in percezione che in produzione.

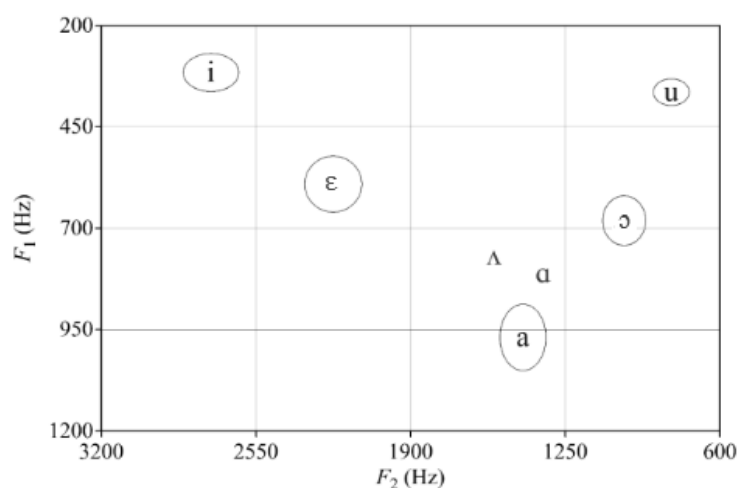
Al fine di osservare in modo più diretto come vengono realizzati i suoni L2, sono stati sviluppati altri approcci come, ad esempio, un approccio combinato tra i *training* di percezione dei suoni L2 e i *training* di produzione (es. Aliaga-Garcia e Mora 2009, Wong 2013), che hanno dimostrato l'effettiva efficacia nel migliorare la produzione dei suoni L2. Ci sono studi, seppur pochi, nei quali l'attenzione è stata orientata esclusivamente verso il ruolo di un *training* di produzione dei suoni L2, dimostrando l'efficacia nella pronuncia dei suoni L2 attraverso alcune misurazioni acustiche (es. Pillot-Loiseau *et al.* 2013) o giudizi da parte di parlanti nativi (es. Akahane-Yamada *et al.* 1998). Un aspetto importante che emerge dagli studi sui *training* di produzione è relativo al *feedback* visivo, che può essere diretto o indiretto e che permette agli apprendenti di confrontare la propria produzione con il suono *target* prodotto da un parlante nativo che funge da modello di riferimento (Kartushina *et al.* 2015). Il *feedback* diretto fornisce una visualizzazione dinamica della posizione e dei movimenti degli articolatori (lingua, labbra ecc.) durante la produzione di un determinato suono. L'ecografo, ossia l'uso di ultrasuoni (Pillot-Loiseau *et al.* 2013; Wilson, Gick 2006), l'elettropalatografo (Schmidt, Beamer 1998) e l'articulografo elettromagnetico (Suemitsu *et al.* 2015) sono tra gli strumenti che possono

essere utilizzati per osservare in modo diretto i movimenti della lingua durante il parlato. Studi come quelli di Pillot-Loiseau *et al.* (2013) e Wilson e Gick (2006) hanno descritto i vantaggi dell'utilizzo di immagini ricavate mediante ultrasuoni per migliorare la pronuncia del contrasto vocalico della lingua francese /y-u/ da parte di apprendenti giapponesi. I risultati mostrano che gli apprendenti che hanno ricevuto un *training* articolatorio con ultrasuono migliorano effettivamente la pronuncia del contrasto non nativo /y-u/, al contrario del gruppo di controllo. Il *feedback* indiretto fornisce, invece, informazioni sulle caratteristiche acustiche dei suoni L2, utilizzando gli spettrogrammi con le prime due o tre formanti (Akahane-Yamada *et al.* 1998) o una rappresentazione grafica astratta che mette in evidenza le principali differenze fonetiche tra i suoni della L1 e quelli della L2. Ad esempio, Carey (2004) ha effettuato un *training* su apprendenti coreani dell'inglese L2 per la produzione delle vocali /æ e ɜ: ə/, fornendo istruzioni verbali relative all'articolazione sia attraverso un video nel quale erano ripresi parlanti nativi e sia con rappresentazioni dello spazio vocalico, ossia con la visualizzazione delle formanti F1 (dimensione verticale correlata all'altezza della lingua) e F2 (dimensione orizzontale correlata all'avanzamento/arretramento della lingua) delle vocali dell'inglese e delle vocali più simili del coreano /ε; o/. I risultati mostrano che gli apprendenti, dopo un *training* di cinque ore, migliorano la produzione di una sola vocale L2, ossia /æ/.

Il *feedback* è, quindi, molto importante per migliorare la pronuncia dei suoni L2 ma, affinché possa essere effettivamente efficace, deve essere facilmente comprensibile e motivante e deve essere fornito subito dopo la produzione L2 (Öster 1997). Inoltre, fornire solo le informazioni acustiche (visualizzazione dei suoni tramite lo spettrogramma e rappresentazioni grafiche su asse cartesiano delle formanti) potrebbe non essere del tutto appropriato per gli apprendenti, dal momento che potrebbero riscontrare delle difficoltà relative a come utilizzare le informazioni acustiche per modificare i propri movimenti articolatori per il raggiungimento di un suono *target* della L2 (Wilson, Gick 2006). Al contrario, approcci che forniscono le informazioni articolatorie possono essere molto più efficaci per gli apprendenti, dal momento che l'informazione articolatoria permette di osservare in modo diretto movimenti della lingua che non potrebbero essere visibili altrimenti (Wilson, Gick 2006). Osservare infatti i video che mostrano i movimenti della lingua dell'insegnante o di un parlante nativo durante il parlato, come anche ricevere il *feedback* insito nel vedere i movimenti della propria lingua in tempo reale può essere estremamente utile (Wilson 2014). Si è osservato inoltre che l'ecografo aiuta a migliorare la pronuncia anche dopo sessioni di *training* molto brevi (es. 30 minuti in Gick *et al.* 2008).

Lo studio pilota descritto in questo articolo riguarda gli effetti che due differenti *training* possono avere sul miglioramento della pronuncia del contrasto vocalico dell'inglese americano /ɑ-ʌ/ (es. *cop-cup*) da parte di apprendenti italofofoni di inglese L2. Gli apprendenti italofofoni mostrano particolare difficoltà nel percepire e produrre questo contrasto (Escudero *et al.* 2011; 2014; Flege *et al.* 1999) poiché, in base all'influenza del sistema fonetico-fonologico della lingua materna, gli italofofoni tendono a realizzare L2-/ɑ/ come L1-/ɔ/ e L2-/ʌ/, spesso, come L1-/a/ (Best, Tyler 2007; Escudero *et al.* 2011; Flege *et al.* 1999). Un altro fattore da considerare è l'ortografia. Nella lingua italiana l'ortografia è trasparente e, per tale motivo, gli italofofoni tendono a pronunciare in modo errato le parole scritte in inglese, la cui ortografia è invece opaca (Bassetti, Atkinson 2015). Ad esempio, gli apprendenti italofofoni tendono a realizzare la parola inglese "cop" con la vocale L1-/ɔ/ per associazione grafema <o> vocale /ɔ/. Lo scopo di questo studio è quello di valutare gli effetti di due *training* differenti, uno percettivo e uno articolatorio con ecografo, o ultrasuono (US), sulla produzione del contrasto non nativo /ɑ-ʌ/. La tecnica di

visualizzazione di immagini tramite ultrasuono è una tecnica non invasiva che permette di osservare la posizione e i movimenti della lingua in tempo reale durante il parlato. Sotto il mento del soggetto viene posizionato un trasduttore (sonda) che contiene dei cristalli piezoelettrici che emettono un fascio di onde sonore ad alta frequenza che raggiunge la lingua e che viene nuovamente riflesso (eco) al trasduttore. Ne risulta, quindi, un segnale riflesso dal quale si genera l'immagine della superficie della lingua (Gick *et al.* 2008). Ai fini del nostro studio, le immagini ricavate con questa tecnica possono essere molto utili poiché i *target* vocalici sono caratterizzati da diverse posizioni della radice e del dorso della lingua, che possono essere visibili con l'ultrasuono (posizione sagittale). Le due vocali, infatti, differiscono sia in altezza (F1) che lungo l'asse antero-posteriore (F2), poiché, rispetto ad /ʌ/, la produzione di /ɑ/ richiede che la radice della lingua sia spostata posteriormente e il dorso sia più basso; al contrario, per produrre /ʌ/ la radice della lingua si sposta anteriormente e il dorso più in alto, come si osserva dalla disposizione delle aree di esistenza acustiche rappresentate nella Figura 1.



**Figura 1:** Spazio vocalico delle vocali della varietà di italiano salentino (ellisse) e dell'inglese americano /ɑ-ʌ/ realizzate da parlanti nativi di sesso femminile. (ellisse= $\pm 1$  SD).

## 2. Obiettivi e ipotesi

Gli obiettivi dello studio sono di osservare: 1) gli effetti di un addestramento esplicito sulla produzione del contrasto L2 /ɑ-ʌ/ da parte di apprendenti italofofoni; e 2) se il *training* articolatorio (effettuato con l'utilizzo di un ecografo) abbia una efficacia maggiore rispetto al *training* percettivo. Entrambi i *training* possono avere effetti a breve termine, inducendo gli apprendenti a spostare il bersaglio articolatorio dal suono più simile della L1 verso un nuovo *target* vocalico. Nonostante ciò, ci aspettiamo che il *training* articolatorio con ultrasuono abbia una maggiore efficacia rispetto a quello percettivo, poiché questo tipo di addestramento si è rivelato utile anche nel caso di addestramenti di breve durata e poiché un *feedback* diretto può contribuire a capire meglio la posizione della lingua per la produzione delle vocali L2 o almeno a differenziare maggiormente la produzione rispetto ai suoni più simili della L1. Inoltre, la pratica con l'ultrasuono permette agli apprendenti un miglior controllo della posizione della lingua, dando la possibilità di riuscire con più facilità, quindi, a raggiungere una posizione articolatoria stabile.

### 3. Metodo

#### 3.1. Soggetti

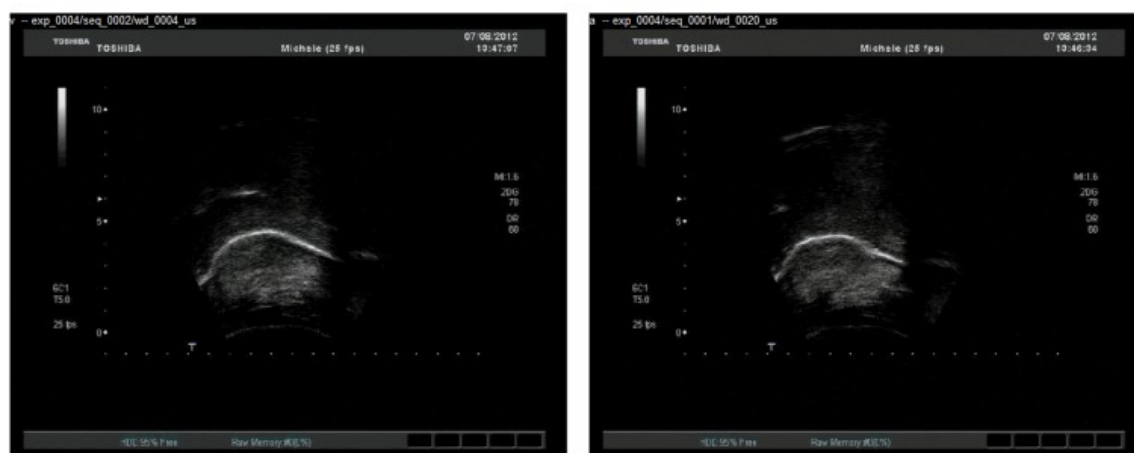
Nove soggetti di sesso femminile (età media: 23; d.s 0.63) hanno partecipato all'esperimento. Tutti i soggetti sono parlanti madrelingua della varietà di italiano di Lecce e, prima di procedere con l'esperimento, hanno compilato una scheda sociolinguistica nella quale veniva richiesto di indicare i periodi trascorsi all'esterno, gli anni di studio della lingua inglese, la lingua madre dei docenti se (italiano o inglese), età e provenienza. Ne risulta che le partecipanti non sono state all'estero per un periodo superiore a un mese, hanno cominciato a studiare la lingua inglese a una età media di 8,5 d.s. 0.52) anni ma sono state esposte principalmente alla lingua inglese parlata da insegnanti italiani. Le partecipanti sono state divise in modo casuale come segue: a) 3 soggetti sperimentali hanno effettuato il *training* articolatorio con ultrasuono (ES-US), SPK1, SPK2 e SPK3; b) 3 soggetti sperimentali hanno effettuato il *training* percettivo (ES-P), SPK4, SPK5 e SPK6; c) 3 soggetti di controllo non hanno ricevuto alcun tipo di *training* (CS), SPK7, SPK8 e SPK9. Inoltre, sono stati raccolti dati relativi a parlanti native americane, ossia dati articolatori registrati con ultrasuono grazie alla collaborazione di una nativa di sesso femminile (AES, 21 anni, Oregon), i cui dati sono stati utilizzati come modello visivo per il gruppo del *training* articolatorio con ultrasuono, e dati acustici di cinque native americane i cui dati sono utilizzati come stimoli per il gruppo del *training* percettivo.

#### 3.2. Stimoli e training

Per i *training* è stata prevista una sessione sperimentale della durata di un'ora. Entrambi i gruppi sperimentali hanno ricevuto informazioni sulla relazione tra pronuncia ed ortografia, ossia circa la corrispondenza grafema-fonema del contrasto L2, e sulla fonetica, ossia circa le differenze fonetiche tra le vocali non native /ɑ-ʌ/ e le vocali più simili in italiano /a-ɔ/. Inoltre, sono state presentate anche le rappresentazioni grafiche della F1 e della F2 (ossia, delle formanti correlate alla posizione della lingua lungo l'asse verticale e l'asse antero-posteriore, rispettivamente), e dello spazio vocalico delle vocali L2 /ɑ-ʌ/ rispetto alle vocali italiane (Figura 1), con istruzioni verbali sulla loro articolazione. Successivamente, i due gruppi sperimentali hanno effettuato la sessione di *training*. Va precisato che le vocali /ɑ-ʌ/ differiscono sia per le caratteristiche spettrali che per la loro durata, poiché /ɑ/ ha una durata maggiore di /ʌ/. Ai fini del nostro esperimento, ci siamo concentrati sulle caratteristiche spettrali e, per tale motivo, gli stimoli sono stati normalizzati per la durata in modo tale da minimizzare il più possibile l'influenza della durata sulla discriminazione delle due vocali non native (Escudero *et al.* 2011).

Ciascun partecipante del gruppo sperimentale ES-US ha effettuato il *training* acustico-articolatorio con Ecografo (Toshiba Aplio XV), che permette di avere un *feedback* immediato grazie alla visualizzazione della posizione della lingua dell'apprendente insieme alla visualizzazione di un video della posizione della lingua del parlante nativo che viene utilizzato, quindi, come modello visivo (Wilson 2014). I dati in produzione del parlante nativo sono stati precedentemente acquisiti e analizzati, e le migliori sequenze di immagini e video sono state selezionate come stimoli da presentare ai partecipanti. Innanzitutto, ai partecipanti è stato mostrato un video (con audio) dei contorni della lingua del parlante nativo durante la produzione di /ɑ-ʌ/ in diversi contesti: in isolamento (Figura 2), nella produzione di parole CVC (es. /bVb/, /pVp/ e /mVsm/, in cui V è /ɑ/ o /ʌ/) e infine nella produzione di una frase completa (es. "I see /bVb/-/pVp/-

/mVm/ inside”). Per chiarire come interpretare i profili linguali e i movimenti della lingua, si è preso in considerazione un gesto di riferimento (Ouni 2011): nella frase “I see /bVb/-/pVp/-/mVm/ inside”, le vocali delle parole “see” e “inside” permettono di avere sia un punto di riferimento iniziale da cui comincia il gesto della vocale bersaglio (spostamento da /i/ di “see” verso la vocale *target* /ʌ/ o /ɑ/) che un riferimento finale in cui il gesto della vocale bersaglio finisce (/i/ di “inside”). In pratica, partendo dalla vocale anteriore alta /i/ (stessa articolazione per inglese e italiano) la lingua deve abbassarsi per il raggiungimento delle vocali /ʌ/ o /ɑ/ per poi innalzarsi di nuovo verso /i/, che pur non facendo parte dell’inventario fonologico italiano è una vocale anteriore alta ma leggermente più aperta di /i/ (Roach 2000). Benché i video fossero accompagnati dall’audio, ai soggetti è stato chiesto di concentrare l’attenzione sulle immagini ecografiche, e pertanto la percezione non è stata oggetto di *training* esplicito. Una volta che i soggetti hanno affermato di aver capito le differenze tra le due vocali *target* della L2, così come tra queste e le vocali più simili della L1, hanno cominciato ad esercitarsi a produrre i suoni della L2 utilizzando la sonda dell’ultrasuono posizionata sotto il proprio mento. Questa procedura ha permesso ai soggetti di osservare il profilo della propria lingua sullo schermo e di effettuare gli aggiustamenti necessari per raggiungere la posizione della lingua della parlante nativa, grazie alla possibilità di controllare la posizione della propria lingua rispetto alle caratteristiche (anche in percezione) del suono realizzato. Durante questa fase, ai partecipanti è stato dato un *feedback*, ossia un *feedback* positivo nel caso in cui avessero raggiunto la posizione ottimale dei *target* L2 o in caso contrario sono stati incoraggiati a riprovare.



**Figura 2:** Immagini Ultrasuono che mostrano la posizione della lingua per la vocale /ʌ/ (a destra) e /ɑ/ (a sinistra) realizzate dalla parlante nativa americana.

I partecipanti del gruppo sperimentale ES-P hanno effettuato un *training* percettivo, ossia un test di identificazione secondo la procedura HVPT (*High Variability Phonetic Training*) già descritta precedentemente, attraverso un’applicazione web appositamente creata in laboratorio (CRIL). Anche in questo caso la durata del *training* è stata di un’ora. Per gli stimoli uditivi sono state considerate parole inglesi /bVb/, /pVp/ e /mVm/ (in cui V corrisponde a /ɑ/ e /ʌ/) pronunciate da cinque parlanti native americane per quattro volte ciascuna. A questi stimoli *target* sono stati anche aggiunti stimoli di controllo, ossia le parole /bVb/, /pVp/ e /mVm/ in cui V è /i/ o /u/. Tutte le vocali sono state normalizzate a una durata di 200ms. Gli stimoli sono stati somministrati uno per volta attraverso l’ascolto tramite cuffia e ai partecipanti è stato chiesto di associare in modo corretto il suono vocalico a uno dei due simboli non ortografici visualizzati sullo

schermo (“^” per /ʌ/, “@” per /a/, “!” per /i/ e “()” per /u/; le opzioni “@-^” sono state presentate per /a/ o /ʌ/ e “!-( )” per /i/ o /u/ - l’ordine dei simboli è casuale). I simboli non ortografici sono stati utilizzati al fine di evitare confusione dovuta alla incorretta associazione tra i suoni e l’ortografia *target* della L2 (Thomson 2011). Nel caso di risposta corretta, i soggetti ricevevano un *feedback* positivo (la parola “Corretto” compariva sullo schermo accompagnata dalla faccia sorridente di Homer Simpson) per poi passare allo stimolo successivo. Al contrario, in caso di risposta errata i soggetti ricevevano un *feedback* negativo (la parola “Sbagliato” compariva sullo schermo accompagnata dalla faccia con espressione di dubbio di Homer Simpson) e ai partecipanti veniva data la possibilità di riascoltare lo stimolo prima di continuare il test con l’ascolto dello stimolo successivo. Lo stimolo successivo è stato presentato con un ritardo di 500ms per un totale di 150 stimoli somministrati per ogni sessione (2 *target* vocalici x 5 parlanti x 3 contesti x 4 ripetizioni + 2 vocali di controllo x 3 parlanti x 3 contesti).

### 3.3. *Acquisizione e analisi dei dati*

Le ipotesi sono testate mettendo a confronto i dati acustici relativi alle produzioni dei partecipanti prima (pre-test) e dopo (post-test) aver effettuato il *training*. Lo studio prevede, quindi, tre fasi: una fase di pre-test, di produzione in L1 e L2 (in ordine casuale tra i soggetti); una fase di *training* di un’ora per i due gruppi sperimentali, ES-US (*training* articolatorio con ultrasuono) e ES-P (*training* percettivo); una fase di post-test, di produzione in L1 e L2. Il gruppo di controllo non ha svolto alcun tipo di *training* per cui i dati di produzioni per il pre-test e post-test sono stati raccolti a distanza di un’ora.

Per i dati di produzione<sup>1</sup>, il corpus L1 è costituito da parole e pseudo parole /pV<sub>1</sub>pV<sub>2</sub>/, in cui V<sub>1</sub> è una delle cinque vocali native /i, ε, a, ə, u/ e V<sub>2</sub> era /i/ o /a/, proposte all’interno di una frase cornice (es. “Dicevi pV<sub>1</sub>pi in su” o “Diceva pV<sub>1</sub>pa a Ken”). Il corpus L2 è costituito da parole /pV<sub>1</sub>p/ dell’inglese americano, in cui V<sub>1</sub> era /a/ o /ʌ/, proposte all’interno di una frase cornice (es. “I see pV<sub>1</sub>p inside”). Tutti i soggetti hanno letto entrambi i corpora per 12 volte. Le produzioni in L1 e L2 dei tre gruppi sono state acquisite nelle sessioni pre- e post-test. Nel dettaglio, le produzioni dei gruppi sperimentali sono state acquisite prima e dopo la sessione *training*, mentre, come anticipato, le produzioni del gruppo di controllo sono state acquisite dopo l’intervallo di un’ora. I dati acustici e articolatori sono stati raccolti simultaneamente in una stanza insonorizzata del laboratorio CRIL. L’acquisizione dei dati articolatori con ecografo è avvenuta grazie all’utilizzo di una sonda convessa (frequenza di 25Hz) posizionata sotto al mento dei soggetti, lungo l’asse sagittale mediano. Al fine di mantenere stabile la sonda sotto il mento e, soprattutto, per ridurre i movimenti della testa, è stato messo a punto un particolare sistema di fissaggio, utilizzando una mentoniera per il fissaggio della testa e un’asta per il fissaggio della sonda. I dati acustici (22050Hz, 16bit) sono stati analizzati mediante PRAAT (Boersma, Weenink 2016) e i valori delle prime due formanti sono state calcolate al 40% nella parte centrale rispetto all’intera durata della vocale (Chládková *et al.* 2011). Successivamente, i valori medi e di deviazione standard della F1 e della F2 per le vocali *target* realizzate nel pre- e post-test sono state rappresentate su un piano cartesiano per ciascun soggetto. I dati sono stati analizzati statisticamente attraverso una serie di t-test a campioni indipendenti (p<0,5) al fine di confrontare: 1) /a-ʌ/ nel pre- vs

<sup>1</sup> Per i dati di produzione, i corpora per la L1 e per la L2 prevedevano parole e pseudo parole nella sequenza /C<sub>1</sub>V<sub>1</sub>C<sub>2</sub>(V<sub>2</sub>)/ in cui C<sub>1</sub> e C<sub>2</sub> erano /p, m, b/. In questa sede ci soffermiamo a descrivere e a presentare i risultati relativi a /p/.

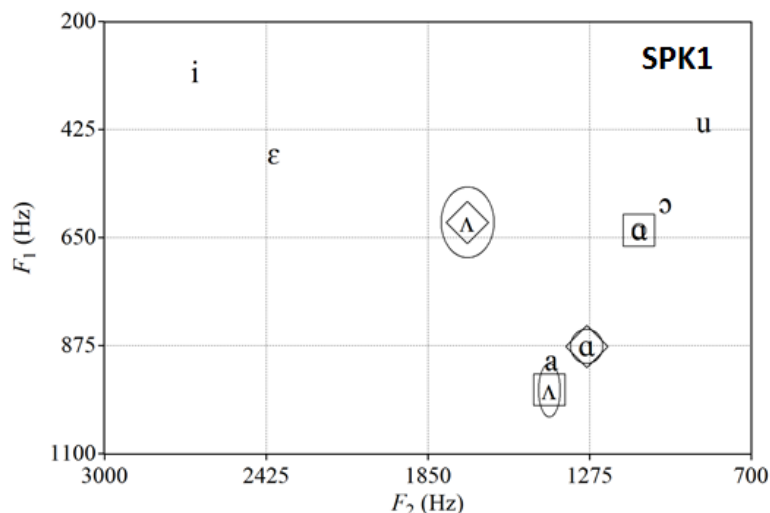


post-test rispettivamente; 2) /a-Λ/ nel pre- e post-test vs le vocali della L1 /a, ə/ rispettivamente. L'efficacia del *training* è stata misurata in base a quanto i *target* della L2 si spostano all'interno dello spazio acustico, per cui quanto più i *target* L2 si spostano dalle vocali della L1 tanto più il *training* sarà stato efficace. Una parte di questa ricerca è stata presentata nello studio (Sisinni *et al.* 2016), in cui sono stati riportati alcuni risultati relativi al solo confronto tra il gruppo sperimentale che ha effettuato il *training* articolatorio e il gruppo di controllo (nessun risultato relativo al gruppo che ha effettuato il *training* percettivo). Nel presente studio, i risultati del gruppo sperimentale che ha effettuato il *training* percettivo sono stati aggiunti e, pertanto, il confronto tra i due tipi di *training* dà un importante contributo al fine di comprendere l'effettiva efficacia di un *training* articolatorio per il miglioramento della pronuncia dei suoni non nativi.

## 4. Risultati

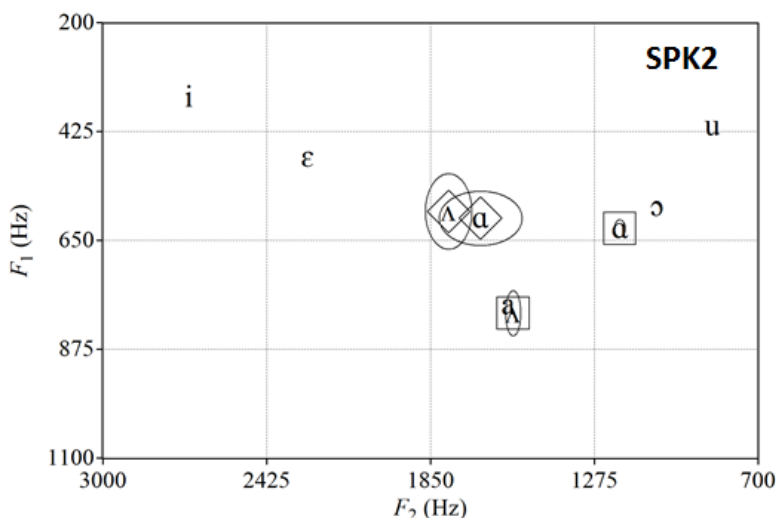
### 4.1. Gruppo training Ultrasuono

La Figura 3 mostra i risultati per le vocali native e non native realizzate dal parlante SPK1. Come si può osservare, nel pre-test la parlante realizza /Λ/ come la vocale nativa /a/, anche se leggermente più bassa [ $t(22)=3,333$   $p<0,05$ ]. Nel post-test, la lingua ha una posizione più alta e avanzata per raggiungere la posizione del nuovo *target* e, infatti, la vocale realizzata differisce in modo significativo sia dalla L1-/a/ che dalla L2-/Λ/ realizzata nel pre-test [ $/Λ-/a/$  F1  $t(23)=-13,132$   $p<0,05$ ; F2  $t(23)=10,006$   $p<0,05$ ;  $/Λ-/Λ/$  pre-test F1  $t(23)=13,207$   $p<0,05$ ; F2  $t(23)=-10,305$   $p<0,05$ ]. La vocale L2-/a/ viene realizzata, nel pre-test, in modo simile alla vocale L1-/ə/, sebbene sia leggermente più bassa e anteriore [F1  $t(22)=6,417$   $p<0,05$ ; F2  $t(22)=9,403$   $p<0,05$ ]. Al contrario, nel post-test il *target* /a/ viene realizzato con una posizione della lingua che differisce in modo significativo rispetto alle vocali L1-/ə/ e L2-/a/ [ $/a-/ə/$  F1  $t(23)=23,763$   $p<0,05$ ; F2  $t(23)=15,752$   $p=0,05$ ;  $/a-/a/$  pre-test F1  $t(23)=-22,540$   $p<0,05$ ; F2  $t(23)=-10,780$   $p<0,05$ ]. Dopo il *training*, quindi, la parlante SPK1 dimostra di essere in grado di realizzare i suoni non nativi riproducendo le loro principali caratteristiche poiché nel post-test la posizione per la vocale L2-/Λ/ è più alta e anteriore rispetto alla vocale nativa /a/, mentre per la vocale L2-/a/ la lingua ha una posizione più bassa e più anteriore della vocale L1-/ə/.



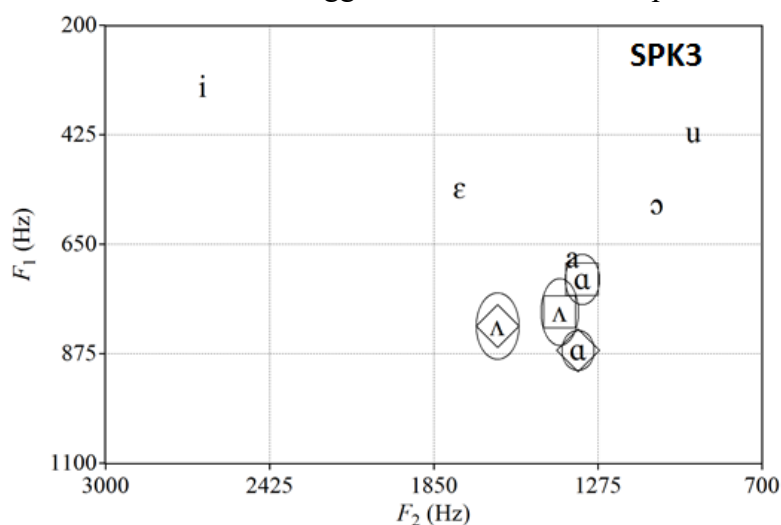
**Figura 3:** Vocali native e non native realizzate da SPK1 nel pre-test (quadrato) e post-test (rombo) (ellisse= $\pm 1$  DS).

Nel pre-test, la parlante SPK2 realizza le vocali L1-/a/ e L2-/ʌ/ senza alcuna differenza, mentre dopo il *training* sia F1 che F2 risultano statisticamente significative in quanto per il raggiungimento del nuovo *target* la lingua assume una posizione più alta e anteriore [/ $\Lambda$ -/a/ F1  $t(23)=-7,592$   $p=0,05$ ; F2  $t(23)=8,281$   $p<0,05$ ; / $\Lambda$ -/ $\Lambda$ / pre-test F1  $t(22)=7,369$   $p<0,05$ ; F2  $t(22)=-9,393$   $p<0,05$ ]. Anche per la parlante SPK2 la vocale L2-/a/ ha caratteristiche simili alla nativa /ɔ/ sebbene risulti essere più bassa e anteriore [F1  $t(21)=4,225$   $p<0,05$ ; F2  $t(22)=8,930$   $p<0,05$ ]. Dopo il *training*, si riscontra una differenza significativa solo per la F2 poiché la vocale L2-/a/ è leggermente più anteriore rispetto alla vocale L1-/ɔ/ [/ $\alpha$ -/ɔ/ F2  $t(22)=14,436$   $p<0,05$ ; / $\alpha$ -/a/ pre-test F2  $t(23)=-11,938$   $p<0,05$ ]. In sostanza, i risultati mostrano che la parlante SPK2 realizza in modo più accurato la vocale non nativa / $\Lambda$ / in quanto riesce a modificare la posizione della lingua sia per la F1 che per la F2 poiché il nuovo *target* risulta essere più alto e anteriore rispetto alla vocale L1-/a/, mentre per la vocale non nativa /a/ si ha uno spostamento significativo solo lungo l'asse antero-posteriore per cui il nuovo *target* ha una posizione più anteriore rispetto alla vocale nativa /ɔ/. I risultati sono rappresentati nella Figura 4 in basso.



**Figura 4:** Vocali native e non native realizzate da SPK2 nel pre-test (quadrato) e post-test (rombo) (ellisse= $\pm 1$  DS).

Al contrario delle due parlanti precedenti, la parlante SPK3 realizza entrambe le vocali non native /ʌ/ e /ɑ/ in modo simile alla vocale nativa /a/ sebbene con una maggiore apertura per entrambe le vocali L2 [F1: /ʌ/-/a/  $t(19)=4,582$   $p<0,05$ ; /ɑ/-/a/  $t(23)=2,392$   $p<0,05$ ]. Nel post-test L2-/ʌ/ differisce in modo significativo sia rispetto alla nativa /a/, con una posizione della lingua più bassa e più avanzata, e sia rispetto alla L2-/ʌ/ del pre-test essendo più anteriore [F1: /ʌ/-/a/  $t(22)=6,095$   $p<0,05$ ; F2  $t(22)=9,264$   $p<0,05$ ; F2 /ʌ/-/ʌ/ pre-test  $t(21)=-7,282$   $p<0,05$ ]. Il fonema L2-/ɑ/ viene realizzato nel post-test con una maggiore apertura sia rispetto alla nativa /a/ che rispetto alla L2-/ɑ/ del post-test [F1: /ɑ/-/a/  $t(23)=12,147$   $p<0,05$ ; /ɑ/-/ɑ/ pre-test  $t(26)=-8,346$   $p<0,05$ ]. In pratica, nel post-test la parlante SPK3 riesce a spostarsi dalla vocale nativa /a/ sebbene entrambe le vocali L2 siano realizzate in uno spazio acustico al di sotto della nativa /a/, ossia con una maggiore apertura per entrambe le vocali e un maggiore avanzamento solo per /ʌ/.



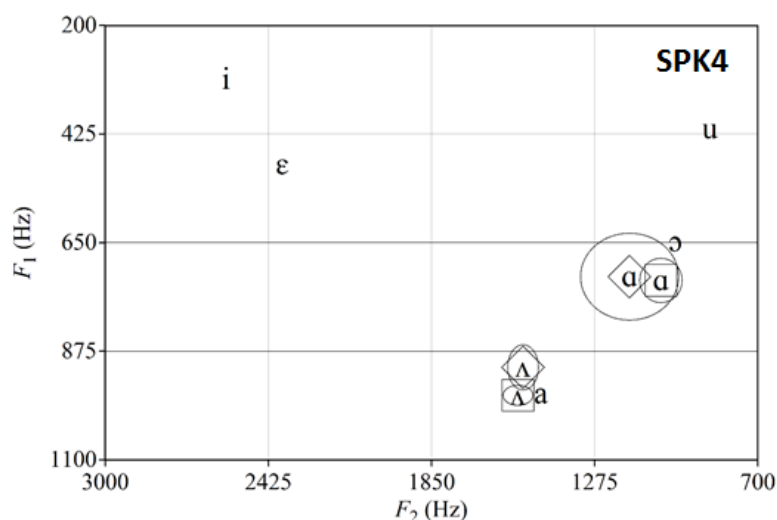
**Figura 5:** Vocali native e non native realizzate da SPK3 nel pre-test (quadrato) e post-test (rombo) (ellisse=±1 DS).

Infine dai risultati del pre-test risulta evidente anche l'influenza dell'ortografia poiché due parlanti su tre (SPK1 e SPK2) associano il grafema <o> al suono della L1 [ɔ]. Solo il parlante SPK3 associa nel pre- e post-test il grafema <o> alla vocale nativa [a] a causa del suo ridotto spazio acustico dovuto al fatto che la parlante realizza [a] con una F1 molto alta e ciò non permette al soggetto di modificare la posizione della lingua in altezza al sopra di [a] perché si avvicinerebbe troppo verso le vocali medio-alte.

#### 4.2. Gruppo training Percettivo

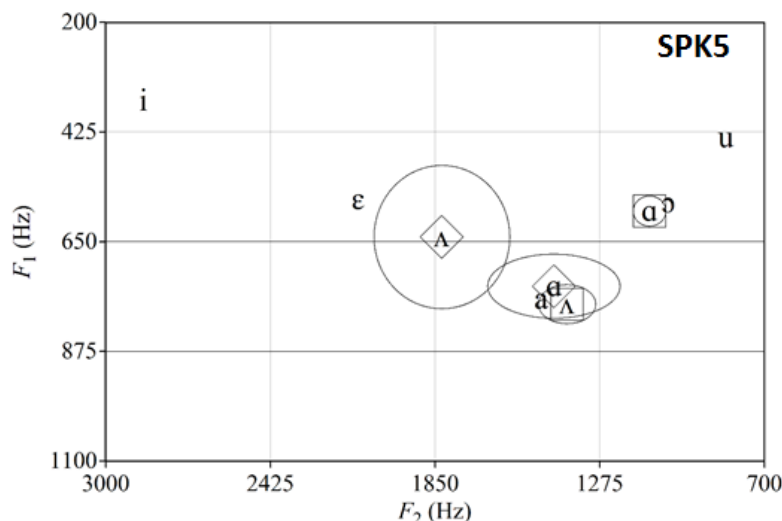
Nel pre-test, SPK4 produce il *target* L2-/ʌ/ in modo simile alla nativa /a/, anche se è significativamente più anteriore [F2  $t(22)=4,618$ ,  $p<0,05$ ]. Nel post-test, al contrario, non si riscontra nessuna differenza significativa tra L2-/ʌ/ e L1-/a/, mentre differisce rispetto a L2-/ʌ/ del pre-test per F1 essendo più alta [F1  $t(20)=3,774$   $p<0,05$ ]. Per quanto riguarda il fonema /ɑ/, nel pre-test viene realizzata come L1-/ɔ/ sebbene con valori di F1 più alti [F1  $t(22)=4,928$   $p<0,05$ ], mentre nel post-test non si riscontra nessuna differenza significativa. Infine, mettendo a confronto /ɑ/ nel pre- e post-test si riscontra una differenza significativa solo per F2 essendo leggermente più anteriore [ $t(25)=-2,066$   $p<0,05$ ]. Ad eccezione di alcune differenze, la parlante SPK4 realizza le vocali non native /ʌ/ e /ɑ/ in modo simile

alle vocali native /a/ e /ɔ/ rispettivamente anche dopo il *training* percettivo. I risultati sono riportati nella Figura 6.



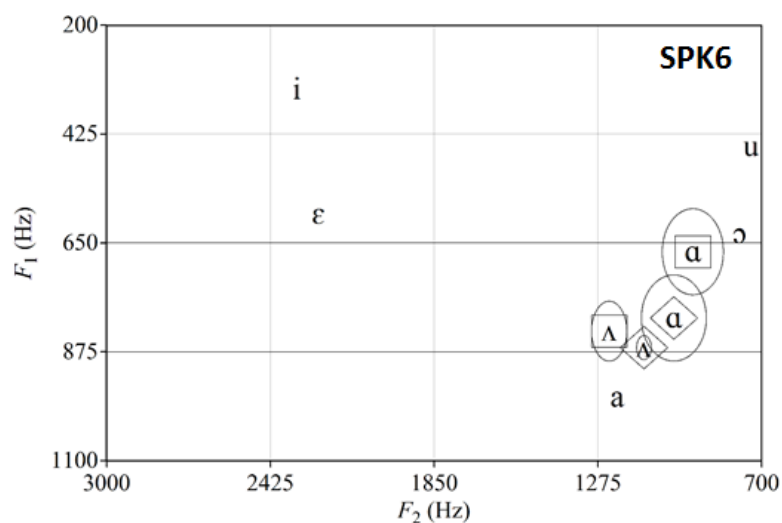
**Figura 6:** Vocali native e non native realizzate da SPK4 nel pre-test (quadrato) e post-test (rombo) (ellisse= $\pm 1$  DS).

Nel pre-test, la parlante SPK5 realizza L2-/Λ/ come la nativa /a/ ma con valori di F2 significativamente più bassi [F2 t(22)=-2,539 p<0,05]. Dopo il *training*, /Λ/ viene realizzata con una posizione più alta e anteriore rispetto alla nativa /a/ [F1 t(14)=-4,070 p<0,05; F2 t(14)=3,943 p<0,05] e, difatti, differisce significativamente anche rispetto alla vocale /Λ/ del pre-test [F1 t(23)=3,230 p<0,05; F2 t(23)=-6,065 p<0,05]. Osservando i valori medi e di deviazione standard si nota che la F1 è di 778Hz (d.s. 40) e di 640Hz (d. s. 147) nel pre- e post-test rispettivamente; mentre la F2 è di 1389Hz (d. s. 100) e di 1826Hz (d. s. 237). I valori di deviazione standard nel post-test sono molto alti e ciò è indice di un ampio grado di variabilità nella produzione delle vocali *target* all'interno dello spazio acustico. Per quanto riguarda la vocale L2-/α/, la sua realizzazione nel pre-test è molto simile alla vocale nativa /ɔ/ anche se significativamente più anteriore [F2 t(21)=2,625 p<0,05]. Nel post-test, /α/ viene realizzata come una vocale più bassa e anteriore e si distingue in modo significativo sia rispetto alla nativa /ɔ/ e sia rispetto a /α/ post-test [L1-/ɔ/: F1 t(13)=4,063 p<0,05; F2 t(13)=2,743 p<0,05; /α/-pre-test: F1 t(21)=-6,973 p<0,05; F2 t(21)=-4,642 p=,005]. Osservando la Figura 7, si può notare che nel post-test /α/ si realizza tra la vocale L1-/a/ e la vocale L2-/Λ/ pre-test. Dal punto di vista statistico, /α/ differisce in modo significativo solo rispetto alla nativa /a/ e solo per la F1 essendo una vocale più chiusa [F1 t(13)=-1,764 p<0,05]. Anche per /α/ i valori di deviazione standard sono molto alti nel post-test. La F1 è di 587Hz (d.s. 31) e 741Hz (d.s. 66) nel pre-test e post-test rispettivamente, mentre la F2 è di 1298Hz (d.s. 23) e 1105Hz (d.s. 231) nel pre-test e post-test rispettivamente. I risultati, quindi, suggeriscono che la parlante SPK5 nel post-test realizza le vocali della L2 spostandosi dalle vocali più simili della L1 ma con un'ampia variabilità (alti valori di deviazione standard), indice di incertezza nella produzione e articolazione delle vocali della L2.



**Figura 7:** Vocali native e non native realizzate da SPK5 nel pre-test (quadrato) e post-test (rombo) (ellisse= $\pm 1$  DS).

La parlante SPK6 realizza /Λ/ nel pre-test con una posizione vicina alla nativa /a/ ma con una F1 significativamente più bassa [F1  $t(21)=-6,468$   $p<0,05$ ]. Dopo il *training* percettivo, si ha una differenza significativa per F1 e F2 poiché la vocale /Λ/ si realizza come una vocale più chiusa e posteriore [F1  $t(12)=-4,450$   $p<0,05$ ; F2  $t(12)=-2,429$   $p<0,05$ ]. Tra pre- e post-test /Λ/ differisce solo per la F2 essendo più posteriore nel post-test [F2  $t(20)=-5,977$   $p<0,05$ ]. Per quanto riguarda /a/, la sua realizzazione è simile alla L1-/ɔ/ anche se la F2 mostra valori significativamente più alti [F2  $t(23)=4,443$   $p<0,05$ ]. Nel post-test, /a/ differisce per entrambe le dimensioni poiché si realizza come una vocale più bassa e più anteriore [F1  $t(14)=7,429$   $p<0,05$ ; F2  $t(14)=4,216$   $p<0,05$ ]. Tra pre- e post-test solo la F1 risulta significativa con una maggiore apertura nel post-test [F2  $t(24)=-3,926$   $p<0,05$ ]. In sostanza, la parlante realizza alcuni cambiamenti nella produzione delle due vocali L2 dopo il *training* ma le principali differenze fonetiche tra /Λ/ e /a/ come anche le loro caratteristiche rispetto alle vocali della L1 /a/ e /ɔ/ non sembrano essere state percepite. Questo si riscontra soprattutto per la F1 poiché nel post-test /Λ/ e /a/ hanno una posizione molto vicina tra loro e tra le vocali L1 /a/ e /ɔ/.



**Figura 8:** Vocali native e non native realizzate da SPK6 nel pre-test (quadrato) e post-test (rombo) (ellisse= $\pm 1$  DS).

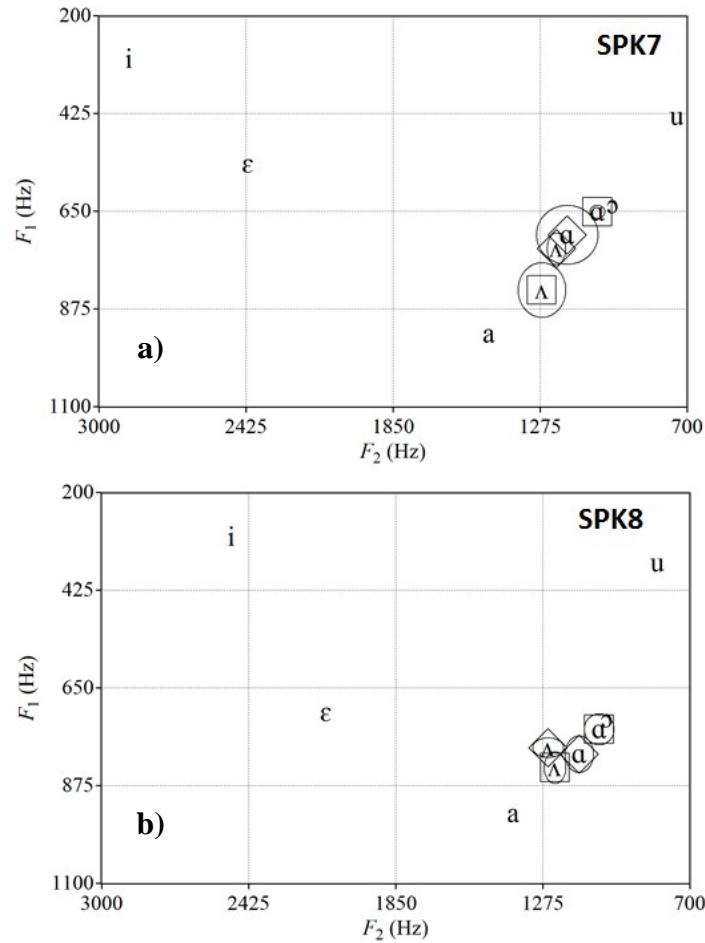
Infine, anche per i soggetti del training percettivo si riscontra un'influenza dell'ortografia poiché nel pre-test le parlanti tendono ad associare il grafema <o> (della parola *pop*) alla vocale L1-/ɔ/, mentre per SPK4 si riscontra anche nel post-test.

### 4.3. Gruppo di controllo

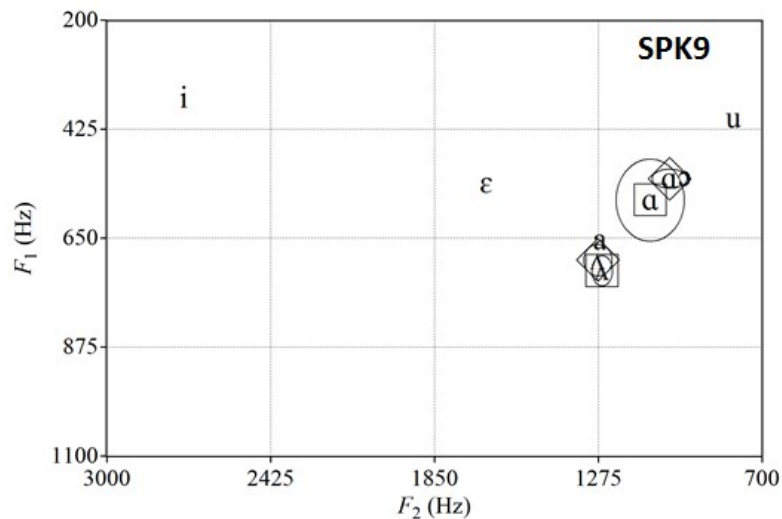
La Figura 9 mostra i grafici F1-F2 per SPK7 (Fig. 9a) e SPK8 (Fig. 9b) del gruppo di controllo. Nel pre-test, entrambe le parlanti distinguono in modo significativo la vocale /ʌ/ per F1 e F2 sia dalla nativa /a/ [SPK7: F1  $t(22)=-4,127$   $p<0,05$ ; F2  $t(22)=-7,162$   $p<0,05$ ; SPK8: F1  $t(19)=-6,437$   $p<0,05$ ; F2  $t(19)=-8,555$   $p<0,05$ ] che dalla vocale /ɔ/ [SPK7: F1  $t(21)=10,144$   $p<0,05$ ; F2  $t(21)=9,344$   $p<0,05$ ; SPK8: F1  $t(20)=8,543$   $p<0,05$ ; F2  $t(20)=11,855$   $p<0,05$ ]. Anche nel post-test, la vocale L2-/ʌ/ viene realizzata tra le due vocali native distinguendosi, quindi, sia per la F1 che per la F2 [ʌ-/a/ SPK7: F1  $t(21)=-9,660$   $p<0,05$ ; F2  $t(21)=-17,480$   $p<0,05$ ; /ʌ-/ɔ/ F1  $t(20)=7,970$   $p<0,05$ ; F2  $t(20)=13,311$   $p<0,05$ ; /ʌ-/a/ SPK8: F1  $t(21)=11,743$   $p<0,05$ ; F2  $t(21)=6,006$   $p<0,05$ ; /ʌ-/ɔ/ F1  $t(22)=6,968$   $p<0,05$ ; F2  $t(22)=11,020$   $p<0,05$ ]. Per la vocale /ʌ/ realizzata nel pre- e post-test si riscontra una differenza significativa solo per F1 che ha valori più alti nel post-test [SPK7: F1  $t(21)=4,257$   $p<0,05$ ; SPK8 F1  $t(20)=3,492$   $p<0,05$ ]. Al contrario, entrambe le parlanti realizzano la vocale non nativa /a/ con una posizione simile alla nativa L1-/ɔ/, anche se per SPK7 è significativamente più anteriore e più bassa [SPK7: F1  $t(21)=2,178$   $p<0,05$ ; F2  $t(21)=4,058$   $p<0,05$ ]. Nel post-test, SPK7 e SPK8 realizzano la vocale L2-/a/ come una vocale più bassa e più anteriore [SPK7: F1  $t(23)=3,182$   $p<0,05$ ; F2  $t(23)=4,627$   $p<0,05$ ; SPK8 F1  $t(22)=5,787$   $p<0,05$ ; F2  $t(22)=5,495$   $p<0,05$ ]. Nonostante alcune differenze tra pre- e post-test nella produzione delle vocali non native, i risultati non indicano il raggiungimento di un nuovo *target*.

Come si può osservare dalla Figura 10 la parlante SPK9 realizza le vocali non native con una posizione simile alle vocali native sia nel pre- che nel post-test. La vocale L2-/ʌ/ viene infatti realizzata con una posizione simile alla nativa /a/ sebbene i valori di F1 siano significativamente più bassi [pre-test:  $t(19)=3,154$   $p<0,05$ ]; post-test:  $t(22)=2,292$   $p<0,05$ ]. La vocale L2-/a/ è significativamente più bassa e anteriore rispetto alla nativa L1-/ɔ/ nel pre-test [F1  $t(26)=2,227$   $p<0,05$ ; F2  $t(26)=3,763$   $p<0,05$ ] e solo leggermente più anteriore nel post-test [F2  $t(22)=2,470$   $p<0,05$ ]. Le produzioni delle vocali non native per la parlante SPK9 non differiscono in modo sostanziale dalle vocali native.

Infine, per quanto riguarda l'influenza dell'ortografia risulta evidente per tutti i soggetti di controllo l'associazione del grafema -o- al suono L1-/ɔ/ nel pre- e nel post-test.



**Figura 9:** Vocali native e non native realizzate da SPK7 (a) e SPK8 (b) nel pre-test (quadrato) e post-test (rombo) (ellisse=±1 DS).



**Figura 10:** Vocali native e non native realizzate da SPK9 nel pre-test (quadrato) e post-test (rombo) (ellisse=±1 DS).

## 4. Discussione e Conclusioni

Il contrasto vocalico dell'inglese americano /ʌ-ɑ/ è particolarmente difficile da produrre per gli apprendenti italofofoni (Flege *et al.* 1999), per via delle caratteristiche del sistema fonetico-fonologico dell'italiano e dell'inglese, nonché dell'ortografia opaca dell'inglese (Bassetti, Atkinson 2015).

In questo studio abbiamo messo a confronto due tipi di *training* - percettivo *vs.* articolatorio – al fine di osservare quale dei due potesse maggiormente contribuire a migliorare la produzione delle vocali non native, come il contrasto dell'inglese americano /ʌ-ɑ/ da noi esaminato. In particolare, i due *training*, della durata di un'ora, messi a confronto sono: 1) un *training* percettivo basato su un test di identificazione secondo la procedura HVPT; e 2) un *training* articolatorio tramite ultrasuono, focalizzato sulla visualizzazione dei contorni e dei movimenti della lingua sia delle produzioni di una parlante nativa (modello di riferimento) in diversi contesti e sia dei movimenti della lingua degli apprendenti coinvolti, con *feedback* in tempo reale. Entrambi i *training* possono contribuire a migliorare la pronuncia dei suoni non nativi, ma sulla base di alcuni lavori discussi in letteratura, si è ipotizzato che l'ultrasuono potesse avere una maggiore efficacia. La nostra ipotesi era, quindi, che il *training* articolatorio potesse dare risultati migliori del *training* percettivo anche nel caso di un addestramento di breve durata, poiché la visualizzazione dei contorni e dei movimenti della lingua può portare i soggetti ad una maggior consapevolezza articolatoria e ad effettuare, quindi, determinati aggiustamenti al fine di raggiungere un nuovo *target* in produzione, distinguendo maggiormente i suoni non nativi da quelli nativi.

In effetti, i risultati confermano la nostra ipotesi principale, a parte mostrare anche la forte influenza dell'ortografia con l'associazione del grafema dell'inglese <o> al suono nativo [ɔ] nel pre-test per tutti i parlanti e nel post-test dai soggetti del gruppo di controllo e da un soggetto del gruppo dei *training* percettivo.

Nel pre-test, indipendentemente dal gruppo, tutti i parlanti realizzano la vocale L2-/ʌ/ come simile alla nativa /a/ e la vocale L2-/ɑ/ come simile alla nativa /ɔ/. Nel post-test, i soggetti del gruppo di controllo, che non hanno ricevuto alcun tipo di *training*, presentano risultati del tutto analoghi a quelli del pre-test, continuando, quindi, a realizzare L2-/ʌ/ come la nativa /a/ e la vocale L2-/ɑ/ come la nativa /ɔ/. Al contrario, alcuni parlanti che hanno effettuato il *training* articolatorio attuano in modo sistematico e significativo strategie per il raggiungimento di una nuova posizione per le vocali non-native. In particolare, due parlanti su tre (SPK1 e SPK2) dopo il *training* articolatorio dimostrano di essere in grado di spostare il bersaglio articolatorio dalle vocali native e di raggiungere una nuova posizione per le vocali della L2 in modo piuttosto accurato. Entrambe le parlanti, infatti, spostano la lingua nella direzione corretta per la vocale L2-/ʌ/, cioè innalzando e avanzando la lingua rispetto alla nativa /a/. Solo SPK1, invece, dimostra di essere in grado di modificare in modo corretto la posizione della lingua anche per la vocale L2-/ɑ/, abbassando la lingua rispetto alla nativa /ɔ/. Per quanto riguarda la terza parlante, SPK3 i risultati sono diversi. In particolare, SPK3 realizza entrambe le vocali non native in uno spazio acustico al di sotto della nativa /a/, probabilmente perché lo spazio acustico delle vocali native /ɛ, a, ɔ/ è troppo ristretto per poter essere modificato, differenziando le vocali non native da quelle native. Per quanto riguarda le parlanti che hanno effettuato il *training* percettivo, solo la parlante SPK5 realizza i *target* L2 in modo differente rispetto alle parlanti native, spostando la lingua nella corretta direzione per entrambe le vocali L2. La vocale L2-/ʌ/ viene realizzata, infatti, con un avanzamento e innalzamento della lingua, mentre la vocale L2-/ɑ/ con un abbassamento e avanzamento



della lingua (anche se la vocale risulta essere troppo vicina alla nativa /a/). Questi cambiamenti mostrano un ampio grado di variabilità (valori alti di deviazione standard) e ciò indica che la parlante realizza le vocali L2 con una incertezza piuttosto che con uno sforzo sistematico e consapevole a produrre i nuovi *target* differenziandoli dalle vocali native come accade, invece, per le parlanti che hanno effettuato il *training* articolatorio. La parlante SPK6 realizza la vocale L2-/ʌ/ spostando la lingua indietro in una posizione vicina alla L2-/ɑ/ e neutralizzando, quindi, le differenze tra le due vocali. Infine, le produzioni della parlante SPK4 rispecchiano quelle dei soggetti di controllo poiché, dopo il *training* percettivo, la parlante realizza i suoni L2 pressoché come le vocali native.

Come ipotizzato, il *training* articolatorio sembra quindi essere particolarmente utile per gli apprendenti al fine di migliorare la produzione del contrasto vocalico non nativo e i suoi effetti sembrano essere più efficaci e stabili rispetto ai risultati ottenuti con il *training* percettivo. Gli apprendenti dimostrano, infatti, di essere in grado di mantenere pressoché la stessa posizione della lingua per i nuovi *target* nel corso delle diverse ripetizioni, come indicato dai valori bassi di deviazione standard. In ogni caso, anche il *training* percettivo comporta un miglioramento nella pronuncia dei suoni non nativi, benché ciò avvenga solo per un parlante su tre (SPK5) e con un ampio grado di variabilità (valori alti di deviazione standard), che mettono in evidenza i diversi tentativi nel produrre i suoni L2. I risultati di questo studio sembrano essere, pertanto, in linea con quanto riportato in letteratura, ossia che entrambi i *training* possono avere un effetto sulla produzione dei suoni non nativi ma il *training* articolatorio sembra dare un contributo più evidente proprio grazie alla visualizzazione e al confronto dei movimenti articolatori dell'apprendente con il modello di riferimento. Inoltre, il *feedback*, soprattutto quello visivo, è estremamente importante, anche per sessioni di *training* brevi come nel nostro caso, perché aiuta ad avere una maggiore consapevolezza dei movimenti degli articolatori per la produzione dei suoni sia della L1 che della L2.

Infine, i risultati di questo studio sono anche in linea con quanto riportato in letteratura per quanto riguarda l'aspetto didattico. Il punto di partenza di tutte le partecipanti, nel pre-test, è stato quello di associare i suoni non nativi ai suoni più simili della L1. Ciò indica che durante il loro percorso di studio dell'inglese come L2 (in media 8.5 anni) le partecipanti non hanno ricevuto, da parte dei docenti italiani di lingua inglese, un insegnamento specifico o efficace sulla pronuncia e sulle differenze tra i suoni della lingua italiana e quelli della lingua inglese. I risultati, quindi, sottolineano l'importanza di fornire agli studenti informazioni focalizzate sulla pronuncia per permettere loro di comprendere le differenze tra le due lingue e di avere una maggiore competenza linguistica attraverso opportune esercitazioni.

Al fine di convalidare i risultati appena descritti sarà necessario incrementare il numero di soggetti ed effettuare *training* di durata superiore e/o sessioni multiple, per far sì che gli apprendenti abbiano un controllo più accurato dei movimenti della lingua al fine di raggiungere una posizione più accurata e precisa del *target* L2. Sarà necessario osservare anche gli effetti a lungo termine, la riorganizzazione del controllo motorio e i suoi effetti a livello fonologico. Inoltre, sarà necessario ampliare il corpus considerando nuovi fonemi, non solo vocalici ma anche consonantici, parole reali inserite all'interno di frasi in contesti comunicativi e domande, per osservare anche gli aspetti suprasegmentali.

**Bionote** Immacolata Sonia d'Apolito si è laureata in Lingue e Letterature straniere nel 2007 ed ha conseguito il Dottorato di Ricerca all'Università del Salento nel 2012 in “Studi linguistici, storico-letterari e interculturali” presentando una tesi dal titolo “La coarticolazione: Studio acustico, cinematico e percettivo di sequenze di sibilanti della lingua francese nelle produzioni di studenti italofofoni”. Attualmente sta frequentando il Dottorato in “Lingue, letterature e culture moderne e classiche”. Collabora presso il CRIL (Centro di Ricerca Interdisciplinare sul Linguaggio). I principali ambiti di interesse sono l'acquisizione di una lingua straniera dal punto fonetico e fonologico da parte di parlanti adulti, soprattutto dal punto di visto articolatorio.

Bianca Sisinni, dopo aver conseguito una laurea in Lingue e Letterature straniere nel 2005, ha conseguito il Dottorato di Ricerca in “Studi letterari, linguistici e culturali” nel 2006. Ha proseguito il suo lavoro di ricerca, pubblicando diversi contributi in sedi nazionali ed internazionali, e ha ricoperto il ruolo di ricercatore di Linguistica Generale, terminato nel 2017. Si è occupata principalmente dell'acquisizione fonetico-fonologica della seconda lingua utilizzando tecniche comportamentali ed elettrofisiologiche. Attualmente è docente di lingua inglese presso la scuola secondaria.

Mirko Grimaldi è Professore Associato di Linguista Generale presso l'Università del Salento. Si occupa di neurobiologia del linguaggio, acquisizione della fonetica e fonologia della L2, e di fonetica e fonologia della variazione dialettale. Dirige il Centro di Ricerca Interdisciplinare sul Linguaggio (CRIL).

Barbara Gili Fivela è Professore Associato presso l'Università del Salento, è presidente dell'Associazione Italiana di Scienze della Voce (AISV, Special-Interest-Group dell'International-Speech-Communication-Association) e vicedirettore del Centro di Ricerca Interdisciplinare sul Linguaggio (CRIL, nel DReAM-Laboratorio di Ricerca Diffusa Applicata alla Medicina). Laureata in Filosofia all'Università di Torino, ha conseguito il dottorato in Linguistica presso la Scuola Normale Superiore (Pisa), specializzandosi nello studio fonetico-fonologico della prosodia e arricchendo la formazione presso l'Ohio State University (U.S.A.), l'Universität des Saarlandes e l'IPDS (Germania), l'LPL (Aix-en-Provence, Francia). Autrice di oltre cento contributi su fonetica e fonologia, attualmente è responsabile di un progetto PRIN sul parlato nella disartria conseguente alla malattia di Parkinson.

**Recapito autori:** [sonia.dapolito@unisalento.it](mailto:sonia.dapolito@unisalento.it); [b.sisinni@gmail.com](mailto:b.sisinni@gmail.com); [mirko.grimaldi@unisalento.it](mailto:mirko.grimaldi@unisalento.it); [barbara.gilifivela@unisalento.it](mailto:barbara.gilifivela@unisalento.it)

## Bibliografia

- Akahane-Yamada R., McDermott E., Adachi T., Kawahara H., Pruitt J.S. 1998, *Computer-based second language production training by using spectrographic representation and HMM-based speech recognition scores*, in *Proceedings of the 5<sup>th</sup> International Conference on Spoken Language Processing (ICSLP 98)*, Sydney, Australia, November 30 – December 4, pp. 1-4.
- Aliaga-Garcia C. and Mora J.C. 2009, *Assessing the effects of phonetic training on L2 sound perception and production*, in *Recent Research in Second Language Phonetics/Phonology: Perception and Production*, Watkins M.A., Rauber A.S. and Baptista B.O. (eds), Cambridge Scholars Publishing, Newcastle (UK), pp. 2-31.
- Barriuso T.A., Hayes-Harb R. 2018, *High Variability Phonetic Training as a Bridge from Research to Practice*, in "The Catesol Journal" 30 [1], pp. 177-194.
- Bassetti B. and Atkinson N. 2015, *Effects of orthographic forms on pronunciation in experienced instructed second language Learners*, in "Orthographic effects in second language phonology. Special Issue. Applied Psycholinguistics" 36 [1], pp. 67-91.
- Best C.T. 1995, *A direct realist view of cross-language speech perception*, in Strange W. (ed.), *Speech perception and linguistic experience: Issues in cross-language research*, York Press, Baltimore, pp. 171-204.
- Best C.T., Tyler M.D. 2007, *Nonnative and second language speech perception: commonalities and complementarities*, in *Second Language Speech Learning: the role of language experience in speech perception and production*, John Benjamins, Amsterdam, pp. 13-34.
- Boersma P. and Weenink D., *Praat: Doing phonetics by computer* [computer program], Version 6.0, 2016. <http://www.praat.org/>
- Bradlow A.R., Pisoni D.B., Akahane-Yamada R., Tohkura Y.I. 1997, *Training Japanese listeners to identify English /r/ and /l/: IV. Some effects of perceptual learning on speech production*, in "The Journal of the Acoustical Society of America" 101 [4], pp. 2299-2310.
- Brown A. 1992, *Survey of Attitudes and Practices Related to Pronunciation Teaching*, AMES, Perth.
- Carey M. 2004, *CALL visual feedback for pronunciation of vowels: Kay Sona-Match*, in "CALICO Journal" 21 [3], pp. 571-601.
- Celce-Murcia M., Brinton M.D., Goodwin J.M. 2007, *Teaching Pronunciation: A Reference for Teachers of English to Speakers of Other Languages*, Cambridge University Press, Cambridge.
- Chládková K., Escudero P., Boersma P. 2011, *Context-Specific acoustic differences between Peruvian and Iberian Spanish vowels*, in "The Journal of the Acoustical Society of America" 130, pp. 416-428.
- Elliot R. 1997, *On the teaching and acquisition of pronunciation within a communicative approach*, in "Hispania" 8 [1], pp. 95-108.
- Escudero P., Benders T., Wanrooij K. 2011, *Enhanced vowel distributions facilitate the learning of second language vowels*, in "The Journal of the Acoustical Society of America" 130 [4], pp. E1206-E1212.
- Escudero P., Sisinni B., Grimaldi M. 2014, *The effect of vowel inventory and acoustic properties in Salento Italian learners of southern British English vowels*, in "The Journal of the Acoustical Society of America" 35 [3], pp. 1577-1584.
- Flege J. E., Mackay I., and Meador D. 1999, *Native Italian speakers' production and perception of English vowels*, in "The Journal of the Acoustical Society of America" 106, pp. 2973-2987.
- Fraser H. (2000). *Coordinating improvements in pronunciation teaching for adult learners of English as a foreign language*, DETI (ANTA. Innovative Project), Camberra.
- Gick B., Bernhardt B.M., Bacsfalvi P., Wilson I. 2008, *Ultrasound imaging applications in second language acquisition*, in Hansen Edwards J.G., Zampini M.J. (eds.), "Phonology and Second Language Acquisition", Ch. 11, John Benjamins, Amsterdam, pp. 309-322.
- Henderson A., Frost D., Tergujeff E., Kautzsch A., Murphy D., Kirkova-Nasckova A., Waniek-Klimczak E., Levey D., Cunningham U., Curnick L. 2012, *English pronunciation teaching in Europe survey: Selected results*, in "Research in Language" 10, pp. 5-27.
- Hismanoglu M., Hismanoglu S. 2010, *Language teachers' preferences of pronunciation teaching techniques: traditional or modern?* in "Procedia Social and Behavioral Sciences" 2, pp. 983-989.
- Iverson P. and Evans B.G. 2009, *Learning English vowels with different first language vowel system II: Auditory training for native Spanish and German speakers*, in "The Journal of the Acoustical Society of America" 126, pp. 866-877.
- Kartushina N., Hervais-Adelman A., Frauenfelder U.H. and Golestani N. 2015, *The effect of phonetic production training with visual feedback on the perception and production of foreign speech sounds*, in "Journal of Acoustical Society of America" 138 [2], pp. 817-832.

- Kirkova-Nasckova A., Tergujeff E., Frost D., Henderson, A., Kautzsch A., Levey D., Murphy D., Waniek-Klimczak E. 2013, *Teachers' views on their professional training and assessment practices: Selected results from the English Pronunciation Teaching in Europe Survey*, in Levis J.M. and LeVelle K. (eds.), *Pronunciation and Assessment: Proceedings of the 4th Annual Pronunciation in Second Language Learning and Teaching Conference*, Iowa State University, pp. 29–42.
- Lee T.S. 2008, *Teaching pronunciation of English using computer assisted learning software: An action research study in an institute of technology in Taiwan*. Doctoral dissertation, Australian Catholic University.
- Lively S.E., Pisoni D.B., Yamada R.A., Tokhura Y. and Yamada T. 1994, *Training Japanese listeners to identify English /r/ and /l/ III. Long-term retention of new phonetic categories*, in “The Journal of the Acoustical Society of America” 96, pp. 2076-2087.
- Logan J.S., Lively S. E., Pisoni D.B. 1991, *Training Japanese listeners to identify English /r/ and /l/: A first report*, in “The Journal of the Acoustical Society of America” 89 [2], pp. 874-886.
- Öster A.-M. 1997, *Auditory and visual feedback in spoken L2 teaching*, in “Reports from the Department of Phonetics”, Umeå° University PHONUM 4, pp. 145–148.
- Ouni S. 2011, *Tongue gestures awareness and pronunciation training*, in *Proceedings of the 12<sup>th</sup> Annual Conference of the International Speech Communication Association, Interspeech 11*, Florence, Italy, August 27-31, pp. 881-884.
- Pillot-Loiseau C., Antolík T.K., Kamiyama T. 2013, *Contribution of ultrasound visualization to improving the production of the French /y/-/u/ contrast by four Japanese learners*, in “Phonetics, Phonology, Languages in Contact”, Paris, France, pp. 86-89.
- Posner M. and Keele S. 1968, *On the genesis of abstract ideas*, in “Journal of Experimental Psychology” 77, pp. 353-363.
- Roach P. 2000, *English phonetics and phonology: a practical course*, Cambridge University Press, Cambridge.
- Schmidt A.M. and Beamer J. 1998, *Electropalatography treatment for training Thai speakers of English*, in “Journal of Clinical Linguistics & Phonetics” 12 [5], pp. 389-403.
- Sisinni B., d’Apolito S., Gili Fivela B., Grimaldi M. 2016, *Ultrasound articulatory training for teaching pronunciation of L2 vowels*, in *Proceedings of the 9th International Conference for Language Learning*, 17-18 November, libreriauniversitaria.it Edizioni, Florence, Italy, pp. 265-270.
- Suemitsu A. and Dang J., Ito T., Tiede M. 2015, *A real-time articulatory visual feedback approach with target presentation for second language pronunciation learning*, in “Journal of Acoustical Society of America” 138 [4], pp. 382-387.
- Thomson R.I. 2011, *Computer Assisted Pronunciation Training: Targeting second language vowel perception improves pronunciation*, in “CALICO Journal” 28 [3], pp. 744-765.
- Wilson S.M., Gick B. 2006, *Ultrasound technology and second language acquisition Research*, in *Proceedings of the 8th Generative Approaches to Second Language Acquisition Conference*, Volume GASLA 2006, Grantham O’Brien M., Shea C. and Archibald J. (eds.), Cascadilla Proceedings Project, Somerville, MA, pp. 148–152.
- Wilson I. 2014, *Using ultrasound for teaching and researching articulation*, in “Acoustical Science & Technology” 35 [6], pp. 285-289.
- Wong J.W.S. 2013, *The effects of perceptual and or productive training on the perception and production of English vowels /l/ and /i:/ by Cantonese ESL learners*, in *Proceedings of Interspeech 14*, pp. 2113–2117.
- Yates L. 2001, *Teaching pronunciation in the AMEP: current practice and professional development*, AMEP Research Centre, <http://www.nceltr.mq.edu.au/conference2001/index.html>.
- Ylinen S., Uther M., Latvala A., Vepsäläinen S., Iverson P., Akahane-Yamada R., Näätänen R. 2010, *Training the Brain to Weight Speech cues Differently: A Study of Finnish second-language users of English*, in “Journal of Cognitive Neuroscience” 22 [6], pp. 1319-32.