

# Guidelines and recommendations for second language pronunciation teaching

## Linee guida e raccomandazioni per l'insegnamento della pronuncia della lingua straniera

---

Federica Cavicchio

Dipartimento di Studi Umanistici, Università del Salento – federica.cavicchio@unisalento.it  
<https://orcid.org/0000-0002-2795-8396>

---

### ABSTRACT

An intelligible speech (i.e., speech that can be easily understood by an interlocutor) is a realistic target for learners of a foreign language, surely more than speech without any accent. This paper reviews the most recent research on the perception and production of both segmental (e.g., speech sounds) and suprasegmental (e.g., accent, rhythm, tone, intonation) characteristics by speakers of a second language (L2) learned in a classroom. Researchers and teachers have suggested numerous ways to apply technology to facilitate the learning of L2 pronunciation. However, many teachers still feel insecure about methods of teaching pronunciation, and the idea of using computers, mobile devices, or other technologies in the classroom can seem sometimes intimidating. In this paper, we will look at technology by focusing on pedagogical tasks, choosing the most effective support tools to achieve the best results in the classroom.

Un parlato intelligibile (cioè un parlato che può essere facilmente comprensibile da un interlocutore) è un obiettivo più realistico per gli studenti di una lingua straniera che un parlato privo di qualsiasi accento. Questo contributo revisiona le ricerche più recenti sulla percezione e produzione delle caratteristiche sia segmentali (es. suoni del linguaggio) che soprasedimentali (es. accento, ritmo, tono, intonazione) da parte di parlanti di una seconda lingua (L2) imparata in classe. Ricercatori e insegnanti hanno suggerito numerosi modi per applicare la tecnologia all'insegnamento della L2 e facilitare l'apprendimento della pronuncia. Tuttavia, molti insegnanti si sentono ancora insicuri sui metodi per insegnare la pronuncia e l'idea di usare computer, dispositivi mobili o altre tecnologie può sembrare doppiamente intimidatorio. Se guardiamo alla tecnologia concentrandoci sui compiti pedagogici e poi sulla scelta degli strumenti più efficaci di supporto per ognuno, possiamo ottenere i migliori risultati sia per gli insegnanti che per gli studenti.

### KEYWORDS

Learning abilities in the 21<sup>st</sup> century, Training practices, Linguistics  
Abilità di apprendimento nel Ventunesimo secolo, Pratiche formative, Linguistica

## ACKNOWLEDGMENTS

Questa ricerca è stata supportata dal progetto REFIN – *Research for Innovation* – Regione Puglia. Titolo del progetto: “Dal cervello alla voce: metodologie e tecniche didattiche innovative per l’apprendimento fonetico-fonologico della seconda lingua in classe”.

## CONFLICTS OF INTEREST

L’Autrice dichiara che non sussistono conflitti d’interesse.

## 1. Introduzione

Insegnanti e ricercatori concordano che la maggior parte degli studenti di seconda lingua (L2) non apprenderanno mai a parlare con lo stesso accento dei parlanti madrelingua (L1) e che parlare anche fluentemente con un accento non nativo è comunque perfettamente accettabile (Derwing & Munro, 2009). Numerosi studi recenti hanno infatti dimostrato che essere comprensibili è un obiettivo più realistico per gli studenti e gli insegnanti di L2 che un parlato completamente privo di accento (Derwing & Munro, 2015). Il principio di intelligibilità guida le ricerche più recenti sulla pronuncia L2 (per una revisione si veda Levis, 2020). Queste ricerche generalmente concordano che le caratteristiche segmentali e suprasegmentali svolgono un ruolo importante sull’intelligibilità e che l’apprendimento esplicito della pronuncia L2 può avere un impatto positivo sulla comprensibilità del parlato L2 (Derwing et al., 1998; Lee et al., 2014). Allo stesso tempo però, sia gli insegnanti che gli studenti di L2 considerano desiderabile parlare L2 con accento nativo (Sifakis & Sougari, 2005; Scales et al., 2006). Il principio di natività (cioè la convinzione che la pronuncia nativa sia realizzabile e desiderabile), è un obiettivo implicito in molte classi L2. Nonostante ciò, studi recenti dimostrano come la formazione della pronuncia in classe sia intermittente, principalmente perché gli insegnanti hanno relativamente poche conoscenze su come insegnare e valutare la pronuncia L2 (Baker e Murphy, 2011; Couper, 2017). Ci troviamo quindi di fronte a una disconnessione tra la pratica pedagogica e risultati delle ricerche sul parlato L2. Questo è dovuto dall’attenzione sporadica dedicata all’insegnamento della pronuncia da un lato, e alla ricerca incentrata sull’intelligibilità del parlato dall’altro. Inoltre, quando la pronuncia è affrontata in classe, gli insegnanti tendono a concentrarsi sulla produzione segmentale probabilmente perché i libri di testo tendono a focalizzarsi sulle caratteristiche segmentali della L2 (Levis, 2016).

Le ricerche sulla pronuncia L2 non hanno uno standard unico per determinare se un determinato tipo di formazione ha successo o meno. Ci sono vari modi in cui viene misurato il miglioramento della pronuncia L2, ad esempio le analisi acustiche, i giudizi di intelligibilità e di comprensibilità e/o di accento straniero da parte di un parlante L1. Ci sono anche vari modi in cui viene insegnata la pronuncia, ad esempio, leggendo e ripetendo singole parole, frasi o paragrafi, o la ripetizione di parlato L1 semi-spontaneo (ad esempio da sit-com o film).

È importante considerare come punto di partenza quali sono i fattori che hanno un ruolo fondamentale sulla pronuncia L2. Il primo è la L1 di uno studente o di una classe e le differenze fonetiche con la L2. Purtroppo, gli studi che indagano su gruppi di studenti L1 che vogliono imparare la stessa L2 riportano spesso risultati contrastanti. Ad esempio, i parlanti giapponesi in Haslam (2011) non mo-

strato miglioramenti nella produzione inglese del contrasto /l/ e // anche dopo il *training* fonetico, mentre altri studi (ad esempio, Hardison, 2003; Hazan et al., 2005) hanno mostrato miglioramenti su questo stesso contrasto tra gli studenti giapponesi. In un altro caso, gli L1 mandarino in Wang (2002), formati alla percezione delle vocali L2 inglesi, non sono migliorati nella produzione, ma gli L1 mandarino in Thomson (2011) hanno invece mostrato un miglioramento. Vista l'incoerenza di alcuni di questi risultati, è auspicabile che i ricercatori che si occupano di pronuncia L2 guardino oltre ai contrasti L1-L2 nella valutazione del successo dell'apprendimento fonetico. Ulteriori fattori che influenzano l'apprendimento fonetico L2 includono l'età di apprendimento degli studenti (Baker, 2011; Aoyama et al., 2008), la qualità delle interazioni in L2 (con insegnanti madrelingua o meno, Derwing & Munro, 2015), i fattori motivazionali (Nagle, 2018) e il coinvolgimento degli studenti nelle attività didattiche (Jenkins, 2004).

Questo contributo considera l'insegnamento della pronuncia L2 da due prospettive: 1) definire le priorità della classe mettendo in primo piano la comprensibilità e l'intelligibilità del parlato L2 e 2) le ricerche formali sul parlato e la pronuncia L2. Infine, suggeriamo alcuni strumenti tecnologici che possono essere usati in classi e al di fuori per migliorare l'apprendimento della pronuncia L2.

## 2. Stabilire le priorità per l'apprendimento fonetico L2

Quando si tratta di determinare quali caratteristiche della pronuncia abbiano bisogno di uno specifico *training*, alcune ricerche hanno dimostrato che alcuni contrasti fonemici L1-L2 sono molto facili, tanto da non aver bisogno di un apprendimento specifico. Ad esempio, gli studenti L1 mandarino di L2 inglese in Derwing et al. (2012) hanno dimostrato di percepire accuratamente l'accento delle frasi, l'intonazione e la distinzione tra /-ti:n/ e /-ti/ anche in assenza di istruzioni specifiche. Sebbene non sia possibile concludere da tali risultati se una percezione accurata si tradurrà in una produzione accurata, è importante sapere quali gruppi di fonemi siano già padroneggiati dagli studenti, in modo che gli insegnanti non perdano troppo tempo su caratteristiche fonetiche che non hanno bisogno di essere particolarmente approfondite.

Al fine di determinare con quali aspetti della pronuncia L2 gli studenti hanno difficoltà e, quindi, quali aspetti dovrebbero essere al centro della formazione in aula, gli insegnanti sono incoraggiati a sviluppare una valutazione della pronuncia come descritto in Derwing e Munro (2015). Gli insegnanti dovrebbero dunque prendere in considerazione la possibilità di raccogliere campioni di parlato degli studenti, sia da letture che da conversazioni, e valutare tali campioni sia globalmente che individualmente in modo da determinare le difficoltà di pronuncia individuali e della classe. Oltre alla produzione linguistica, un compito percettivo che richieda agli studenti di dimostrare la loro capacità di percepire le distinzioni segmentali e soprasedimentali della L2 può guidare ulteriormente lo sviluppo di un curriculum specifico per l'apprendimento della pronuncia. Sulla base dei risultati della valutazione percettiva, gli insegnanti sono in grado di stabilire quali sono i fonemi L2 che causano maggiore difficoltà e influenzano l'intelligibilità del parlato della classe e che dovrebbero quindi essere al centro dell'apprendimento fonetico. Un altro parametro di valutazione è il carico funzionale che, a livello segmentale, può essere determinato in base al numero di coppie minime. Ad esempio, per l'inglese il contrasto tra /l/ e /n/ distingue più parole del contrasto /d/ e /ð/ (Munro e Derwing, 2006). Chiaramente, per L1 italiani, il contrasto /l/-/n/ non è

particolarmente impegnativo, mentre il contrasto /d/-/ð/ lo è. In conclusione, per stabilire le priorità di pronuncia L2 sia segmentali che suprasegmentali, gli insegnanti dovrebbero valutare la produzione e percezione dei suoni L2 dei loro studenti, e considerare fino a che punto la produzione di determinate distinzioni gioca un ruolo nell'intelligibilità del parlato degli studenti.

### 3. Valutazione dell'efficacia dell'apprendimento fonetico

Gli studenti di L2, specialmente quelli nelle prime fasi dell'apprendimento, tendono a mostrare miglioramenti nella loro pronuncia nel tempo. Pertanto, al fine di determinare se un determinato tipo di apprendimento fonetico è efficace, è importante fare riferimento a ricerche che includono sia un gruppo che riceve un diverso tipo di *training* fonetico, sia un gruppo di controllo che non riceve alcun *training*. Inoltre, un post-test ritardato consente ai ricercatori di determinare se gli effetti del *training* sono di lunga durata o meno (Thomson & Derwing, 2015).

Il miglioramento della pronuncia può essere determinato principalmente in due modi: tramite la valutazione degli ascoltatori e le analisi acustiche. Mentre le valutazioni di comprensione degli ascoltatori da parte di parlanti L1 sono considerate il *gold standard* nelle ricerche sulla pronuncia (Derwing & Munro, 2009), alcuni studi di fanno anche uso di analisi acustiche. Gran parte delle ricerche che indagano l'efficacia dell'apprendimento fonetico utilizzano misure di comprensione quali le valutazioni di comprensibilità (ad esempio, Foote & McDonough, 2017; Martin, 2018) o intelligibilità (ad esempio, Derwing et al., 2014), spesso insieme a valutazioni di fluenza e/o valutazione di accento straniero. Anche il segnale acustico, analizzato manualmente (ad esempio, Counselman, 2015) o automaticamente (ad esempio, Suemitsu et al., 2015; Tejedor-García et al., 2020) sono comuni e possono essere utilizzate per determinare come determinate caratteristiche di pronuncia cambino nel corso del tempo.

Sebbene pochi insegnanti siano in grado di svolgere analisi sistematiche dello sviluppo della pronuncia dei loro studenti, può essere incoraggiato l'uso di metodi di apprendimento della pronuncia la cui efficacia è stata dimostrata attraverso ricerche sul parlato L2. Alcuni di questi lavori sono descritti di seguito. Sfortunatamente, la maggior parte delle ricerche attualmente è concentrata sull'apprendimento dell'inglese come L2. Questa sproporzione si riflette inevitabilmente sulla revisione che segue.

### 4. Ricerche sull'apprendimento della pronuncia L2

Sebbene lo status linguistico di un insegnante L2 (madrelingua o meno) non giochi un ruolo preponderante sulla pronuncia finale degli studenti (Levis et al., 2016), alcune ricerche hanno generalmente dimostrato che istruzioni esplicite, incentrate sull'insegnamento della forma corretta della pronuncia, insieme a istruzioni correttive tramite *feedback* dell'insegnante, forniscono i migliori risultati in termini di apprendimento fonetico L2 (Saito, 2013; Saito & Lyster, 2012).

Un tipo di apprendimento fonetico L2 altamente testato e dimostratosi efficace per migliorare la produzione segmentale L2 attraverso la percezione uditiva è il *training* fonetico ad alta variabilità (*High Variability Phonetic Training*-HVPT), che allena la percezione degli ascoltatori con una quantità relativamente ampia di campioni di vocali e consonanti L2 prodotti da più parlanti madrelingua (per una

revisione, si veda Thomson, 2018). I risultati degli studi HVPT dimostrano un miglioramento delle vocali L2 inglesi da parte di L1 mandarino (Thomson, 2011), greco (Lengeris, 2018), francese (Iverson et al., 2011) e spagnolo (Carlet, 2017), e il miglioramento delle consonanti L2 inglesi in studenti coreani (ad esempio, Huensch & Tremblay, 2015), giapponesi (Bradlow et al., 1997), francesi (Melnik & Peperkamp, 2021) e spagnoli (Carlet, 2017).

Un numero relativamente elevato di ricerche recenti ha inoltre studiato l'efficacia di *training* della pronuncia al di fuori della classe. Queste ricerche hanno sottolineato una serie di vantaggi per gli studenti coinvolti in *training* alla pronuncia assistita da computer (*Computer-Assisted Pronunciation Training*, o CAPT). I vantaggi includono tempo di pratica illimitato e flessibilità di accesso, nonché l'opportunità di essere esposti a input fonetici variati e la possibilità di ottenere una *feedback* immediato sulla propria pronuncia (Engwall et al., 2004; Levis, 2007). Al CAPT viene spesso aggiunto anche un elemento di gaming. Ad esempio, Barcomb e Cardoso (2020) dimostrano l'efficacia del *training* della pronuncia ludicizzato (cioè un *training* che include elementi di un gioco ma che in realtà non è un gioco). Gli studenti di inglese delle scuole medie inferiori giapponesi in quello studio sono stati premiati con punti e badge mentre completavano una serie di attività di percezione e pronuncia incentrate sul contrasto inglese /l/ - /l/. Gli studenti hanno dimostrato una migliore precisione della pronuncia dei suoni addestrati che migliorava nel tempo.

Sebbene l'efficacia dei compiti di ascolto e ripetizione sulla pronuncia possa essere limitata (O'Brien, 2019), un modo molto popolare usato nelle classi L2 per allenare la pronuncia è quello di ascoltare un parlante L1 e poi ripetere il parlato registrando il risultato. Questo compito è comunemente chiamato *shadowing*. Il compito incoraggia gli studenti a concentrarsi anche sugli aspetti soprasegmentali del parlato oltre che sugli aspetti solamente segmentali come nelle ricerche precedenti. Lo *shadowing*, infatti, ha dimostrato di avere vantaggi sia per le caratteristiche segmentali (Zajc & Rojczyk, 2014) che soprasegmentali (Lima, 2015; Foote & McDonough, 2017) del parlato L2. Gli studenti di inglese L2 in Foote e McDonough (2017) hanno completato otto settimane di *shadowing* in cui hanno ripetuto e registrato i dialoghi di una popolare sitcom. Un gruppo di ascoltatori L1 hanno poi valutato le registrazioni e le hanno valutate in termini di comprensibilità, accento e fluenza del parlato. Foote e McDonough riportano che gli studenti avevano atteggiamenti positivi nei confronti dell'attività di *shadowing* e che i giudizi dei parlanti L1 sulla comprensibilità e la fluenza degli studenti miglioravano nel tempo.

#### 4.1 Visualizzazione del parlato L2

Le tecniche di visualizzazione come, ad esempio, la visualizzazione di forme d'onda e spettrogrammi, o immagini a ultrasuoni dei processi articolatori, consentono agli studenti di ricevere *feedback* in tempo reale sul loro parlato L2. Numerose ricerche hanno dimostrato che insegnare agli studenti come interpretare le frequenze delle formanti del parlato L2 utilizzando per la visualizzazione strumenti quali Praat (Boersma & Weenink, 2022) e Audacity (Audacity Team, 2021) può consentire loro di migliorare la produzione delle vocali L2 (Offerman e Olson, 2016; Olson, 2019; Olson & Offerman, 2020; Suemitsu et al., 2015). Purtroppo, le schermate in Praat o altri programmi informatici non sono facili da capire e interpretare per gli insegnanti e gli studenti. Tuttavia, un certo numero di

ricerche sostiene che l'uso di forme d'onda e spettrogrammi può essere utilizzato per l'insegnamento degli aspetti soprasegmentali della L2, in particolare per quel che riguarda gli aspetti di durata delle vocali e dell'intonazione (Chun, 2013; Hardison, 2004; Levis, 1999). Levis e Pickering (2004) hanno dimostrato l'efficacia dell'insegnamento dell'intonazione agli studenti di inglese L2 tracciando i contorni intonativi. Gli studenti di giapponese L2 in Okuno e Hardison (2016) hanno ricevuto un *training* audiovisivo composto da file audio e visualizzazioni di forme d'onda, un *training* solo audio o nessun *training* sulla durata delle vocali in giapponese. I partecipanti ai gruppi sperimentali hanno mostrato che i partecipanti al gruppo audiovisivo hanno migliorato le loro produzioni più dei partecipanti del gruppo solo audio. Allo stesso modo, Motohashi-Saigo e Hardison (2009) hanno dimostrato l'efficacia della visualizzazione delle formanti nell'apprendimento della lunghezza delle vocali giapponesi. Chun e colleghi (2015) hanno infine mostrato che gli studenti di mandarino L2 che hanno confrontato i contorni intonativi della propria produzione con quelli di madrelingua mandarino hanno migliorato la loro produzione di toni.

#### **4.2 Pratiche per migliorare la pronuncia al di fuori della classe**

La tecnologia del riconoscimento vocale automatico (Automatic Speech Recognition-ASR) è progettata per riconoscere le parole pronunciate ed è spesso accompagnata dalla trascrizione in testo scritto, un processo noto anche come "speech-to-text". Chiunque abbia chiesto informazioni a Siri, Alexa o l'assistente vocale Google ha sperimentato l'ASR. L'uso di sistemi di sintesi vocale ha mostrato risultati positivi per il miglioramento della pronuncia. Liakin et al. (2017) hanno scoperto che studenti di francese L2 che hanno utilizzato una applicazione di sintesi vocale sui loro dispositivi mobili hanno migliorato la loro pronuncia in modo simile agli studenti impegnati in conversazioni L2 con gli insegnanti. L'ASR si basa su una combinazione di analisi acustiche e intelligenza artificiale, ed è ampiamente pubblicizzato come un metodo promettente per valutare e fornire *feedback* automatico sulla pronuncia (O'Brien et al., 2018). Purtroppo, sono ancora troppo pochi gli studi che allineano il rilevamento degli errori ASR ai giudizi umani su L2 (ad esempio, Chun, 2013; McCrocklin & Edalatishams, 2020). Molti insegnanti e ricercatori hanno pensato di utilizzare programmi di speech-to-text per l'insegnamento della pronuncia. Il presupposto è che se la trascrizione automatica del parlato di uno studente è per lo più accurata, la pronuncia dello studente è accettabile. Al contrario, se molte parole vengono trascritte in modo errato, allora deve essere presente un problema di pronuncia. La ricerca ha mostrato risultati positivi nell'uso dell'ASR per la pratica della pronuncia indipendente con professionisti adulti (Hincks, 2003) e con i bambini (Mittra et al., 2003; Neri et al., 2008). Quando gli studenti notano le parole dettate non sono state trascritte come previsto, diventano consapevoli dei loro punti deboli e modificano la loro pronuncia fino a quando le loro parole non vengono trascritte in modo più accurato.

#### **4.3 Ulteriori fattori che influenzano la pronuncia L2**

Sebbene sia dimostrato che gli studenti hanno difficoltà a valutare la propria pronuncia (vedi ad esempio, Trofimovich et al., 2016), la consapevolezza degli studenti sulla propria pronuncia può essere positivamente correlata alle valutazioni di



comprensibilità del loro parlato (Kennedy & Trofimovich, 2010). Compiti espliciti che incoraggiano la consapevolezza della pronuncia possono essere quindi particolarmente utili. Ad esempio, Añorga e Benander (2015) e Martin (2018) hanno dimostrato l'efficacia di compiti che incoraggiano gli studenti a confrontare le proprie produzioni con quelle di parlanti L1.

Ulteriori fattori che possono svolgere un ruolo sul miglioramento della pronuncia includono i livelli di competenza linguistica degli studenti, la durata dell'apprendimento e il numero di fonemi addestrati (Sakai & Moorman, 2018). Secondo questo studio, gli studenti con livelli di competenza più bassi tendono a fare progressi più rapidamente rispetto agli studenti più avanzati, e il numero di fonemi mirati dovrebbe essere limitato, possibilmente a un minimo di tre, per ottenere i migliori risultati.

## 5. Tecnologie che facilitano l'apprendimento

Ad oggi sono a disposizione una vasta gamma di strumenti tecnologici per l'apprendimento della pronuncia L2, sia gratuiti che a pagamento. Il numero e la varietà delle tecnologie disponibili possono sembrare scoraggianti per gli insegnanti che sono semplicemente alla ricerca di strumenti efficaci da utilizzare nelle loro classi. Di seguito, elenchiamo alcune domande che gli insegnanti dovrebbero porsi per scegliere lo strumento tecnologico più adatto ai bisogni di apprendimento della classe:

1. adeguatezza agli obiettivi di apprendimento: lo strumento supporta l'insegnante nello svolgimento di compiti didattici specifici (ad esempio *shadowing*)?
2. Qualità e accuratezza: lo strumento fornisce informazioni sull'intonazione accurate?
3. Praticità d'uso: lo strumento è facile da imparare e da usare? Funziona in modo affidabile?
4. Costo: lo strumento è gratuito o economico?

Gli strumenti di illustrati nei prossimi paragrafi soddisfano complessivamente i criteri elencati. Purtroppo, come la maggior parte delle ricerche sulla pronuncia L2, anche la maggior parte degli strumenti sono sviluppati per l'apprendimento della lingua inglese. Chiaramente, quindi, quello che segue non vuole essere un elenco esaustivo di tutti gli strumenti tecnologici disponibili ad oggi, ma una selezione di alcuni strumenti che si sono rivelati utili e affidabili nel tempo. Inoltre, non sono stati inclusi programmi, applicazioni o siti web progettati come corso completo di L2, ma solo le applicazioni specifiche per la pronuncia. Inevitabilmente, è necessario riconoscere che la tecnologia cambia molto velocemente e, al momento in cui l'articolo è scritto, potrebbero esserci aggiornamenti sostanziali agli strumenti tecnologici elencati, o potrebbero essere apparsi nuovi, validi strumenti.

### 5.1 Strumenti utili per lo shadowing

Per quanto utile possa essere ascoltare suoni e parole individuali, gli studenti hanno anche bisogno di un modello di parlato L2 in un contesto verbale allargato.

Tale contesto può essere trovato in molte fonti online, tra cui TED Talks (<http://www.ted.com>), National Public Radio (<http://www.npr.org>), British Broadcasting Corporation (<http://bbc.co.uk>), YouTube (<http://youtube.com>) e altri. Questi materiali raggiungibili gratuitamente online consentono agli studenti di ascoltare più modelli di parlato, il che può aumentare la flessibilità nella comprensione di molte varietà di inglese (Levis, 2007). I video inoltre forniscono materiale utile per lo *shadowing*.

Gli insegnanti che utilizzano video online per lo *shadowing* dovranno fare attenzione che:

1. Le registrazioni includano fonemi, accento delle parole, schemi di intonazione, ecc. su cui si vuol fare pratica
2. La pronuncia di chi parla è chiara e presentata a una velocità ragionevole
3. Il vocabolario usato è a un livello che gli studenti capiscono
4. L'argomento del video è appropriato per gli studenti e per lo scopo di apprendimento. È anche utile sapere se il sito Web di origine consente di riprodurre la registrazione a una velocità inferiore e se è disponibile una trascrizione. In caso contrario, la registrazione può sicuramente essere utilizzata con successo, ma queste funzionalità rappresentano sicuramente un vantaggio.

## 5.2 Sound of Speech

Il sito web più noto per la dimostrazione della pronuncia dei suoni è il sito dell'Università dell'Iowa <http://soundsof speech.uiowa.edu/>. Un'app correlata, Sounds of Speech, è disponibile per dispositivi Apple e Android 2018 e offre funzioni simili. Sounds of speech presenta animazioni e video che descrivono le caratteristiche principali per l'articolazione di vocali e consonanti dell'inglese americano, il tedesco e lo spagnolo. Sia il sito Web che l'app includono un diagramma di sezione sagittale animato per ciascun fonema, una descrizione scritta di come viene prodotto il suono, un breve video di un parlante che pronuncia il suono e parole esempio. La figura 1 mostra la pronuncia di /n/ spagnola dal sito Web Sounds of Speech.

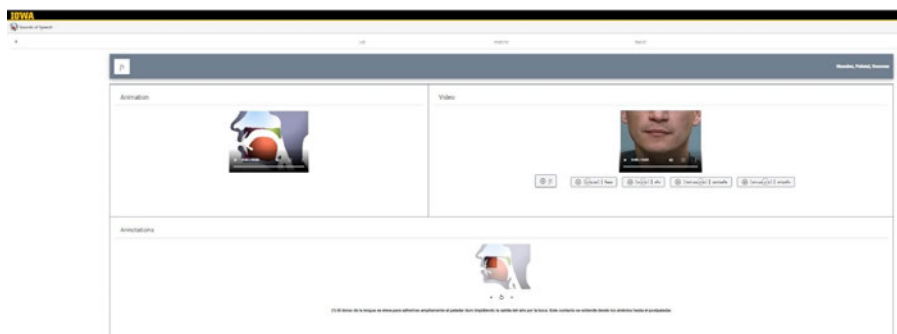


Figura 1. Esempio di schermata da Sound of speech per il fonema /n/ spagnolo.



### 5.3 The Phonetics

The Phonetics è un'app creata da un gruppo di ricerca dell'Università di Lingue Straniere di Tokyo per l'apprendimento fonetico dell'inglese che include illustrazioni animate del sistema articolatorio tridimensionali e mostra la pronuncia di ciascun fonema insieme a parole di esempio registrate. L'app non include video di un vero parlante, come fa invece Sounds of Speech, ma sono incluse animazioni complete per un'ampia gamma di parole campione e coppie minime. La figura 2 mostra la pronuncia di /d/ in The Phonetics. Inoltre, gli utenti possono scegliere di ascoltare la voce di un uomo o di una donna e le registrazioni possono essere riprodotte a velocità normale o più lenta. The Phonetics ed è disponibile solo per i dispositivi Apple.

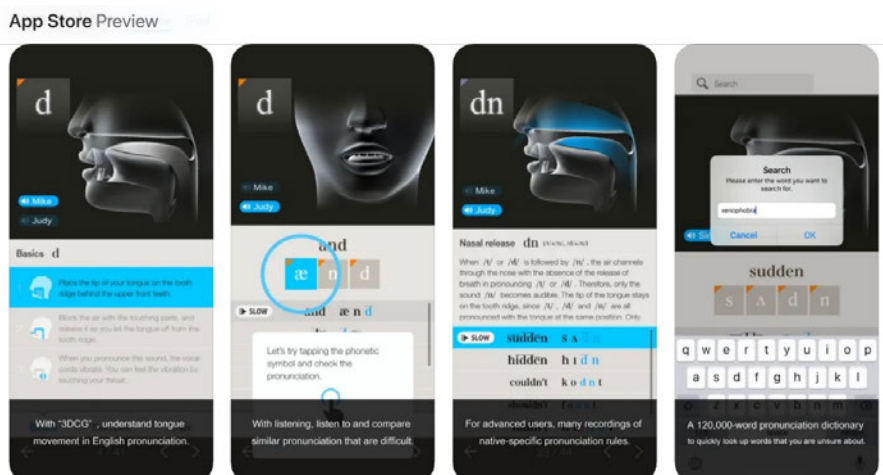


Figura 2: Schermata della preview dell'app *The Phonetics*.

### 5.4 Youglish

Un'ulteriore risorsa utile che colma il divario tra le singole parole e il parlato naturale è YouGlish (<http://youglish.com/>). Questo sito Web gratuito è una sorta di dizionario di pronuncia contestualizzato. Il programma ricerca su video di YouTube esempi di pronuncia sia di parole che di frasi in contesti naturali. Gli utenti possono cercare una parola o una frase e YouGlish trova una selezione di video che la contengono. I video sono impostati per iniziare la riproduzione leggermente prima che appaia la parola scelta e gli utenti hanno la possibilità di ascoltare solo il contesto immediato o l'intero video. Vengono inoltre visualizzati i sottotitoli e i video possono essere riprodotti a velocità normale, più veloce o più lenta. Gli utenti possono specificare se desiderano cercare esempi in inglese americano, britannico o australiano.

### 5.5 Pratica di pronuncia tramite app

Molte app e programmi affermano di analizzare la pronuncia dell'utente e valutarne l'accuratezza. Tuttavia, la maggior parte di questi strumenti danno giudizi

imprecisi sulla pronuncia degli utenti (Derwing & Munro, 2015). Molti studi hanno valutato l'impatto di programmi e applicazioni per smartphone, quali *Windows Speech Recognition* (McCrocklin, 2016), *iFlytek Voice Input* (Liu et al., 2019), *Oral English Drill & Test* (Yuan e Li, 2020), *Fluent English learning* (Liu et al., 2019) ed *Elsa speak* (Kholis, 2021) sull'apprendimento della pronuncia della L2. Altri studi hanno esaminato l'uso dell'ASR nell'insegnamento e nell'apprendimento della pronuncia L2 utilizzando applicazioni ASR sul Web quali *Google Web Speech* (Wallace, 2015) e *Speechnotes* (Evers e Chen, 2020). Quando questo articolo è stato scritto, una ricerca su Google con le parole chiave *Google Web Speech* ha portato a una pagina chiamata *Speech to Text by Google*, che offre una prova gratuita del servizio e una demo. Kaiser (2017) ha testato trenta app per esercitarsi nella pronuncia L2. Molte delle app sono gratuite o molto economiche oppure hanno versioni sia gratuite che a pagamento con funzionalità aggiuntive. Kaiser conclude che non esiste una "killer app" che soddisfi le esigenze di pronuncia di ogni insegnante o studente. Invece di cercare un'app perfetta, sarebbe quindi più produttivo chiedersi quale app potrebbe essere integrata nell'insegnamento della pronuncia sia in classe che fuori. Inoltre, la maggior parte delle app si concentra su fonemi e parole individuali e spesso non presta attenzione alla componente suprasegmentale.

Kaiser (2017) suggerisce agli insegnanti di porsi queste domande prima di utilizzare un'app o un sito Web ASR:

1. L'app risponde agli obiettivi dello studente?
2. Il materiale è accurato e pedagogicamente valido?
3. Chi ha sviluppato l'app e quali credenziali possiede? (Vedasi il sito web degli sviluppatori)
4. Ci sono recensioni o raccomandazioni da fonti autorevoli?
5. Qual è il costo iniziale? Ci sono costi di abbonamenti o in-app nascosti?

## 6. Conclusioni

La recensione della letteratura di settore condotta in questo studio testimonia la facilità di accesso agli strumenti per migliorare la pronuncia L2. Grazie al moltiplicarsi di questi strumenti lo studente di lingue straniere ha facile accesso a una moltitudine di siti web e app che promettono di ridurre l'accento straniero in modo rapido e semplice. Questo articolo è una sintesi di ricerca critica che vuole fornire ai lettori un *background* nel campo della formazione sulla pronuncia L2, assieme ad alcune linee guida e raccomandazioni su come scegliere gli strumenti tecnologici più adatti alle esigenze di apprendimento degli studenti. Nell'insegnamento della pronuncia L2 è importante distinguere quali sono gli obiettivi degli insegnanti e degli studenti, le aspettative dei parlanti in termini di accento straniero e intelligibilità, la pronuncia dei singoli fonemi o dell'intonazione della frase, tenendo a mente quali tipi di *training* e strumenti tecnologici si sono dimostrati più efficaci per raggiungere gli scopi prefissati. Ciò è particolarmente vero se consideriamo l'obiettivo finale dell'apprendimento della pronuncia: un parlato L2 comprensibile. Questa revisione critica chiarisce inoltre perché pochissimi insegnanti abbiano finora sentito parlare di approcci empiricamente guidati all'insegnamento della pronuncia: studi sui metodi per migliorare la pronuncia tramite training specifici sono in gran parte pubblicati su riviste di orientamento tecnico; pochi purtroppo sono accessibili agli insegnanti di lingue. È perciò necessario colmare il divario tra ricerca e pratica, e ciò richiede una migliore cooperazione tra

ricercatori e insegnanti e tra ricercatori e programmatori informatici. Richiede inoltre la disponibilità da parte dei ricercatori di acquisizione del linguaggio e delle lingue seconde di pubblicare, possibilmente in formato *open access*, per gli insegnanti. Fortunatamente la tecnologia può colmare rapidamente il divario tra ricerca e pratiche di insegnamento. Infatti, data la rapidità con cui la tecnologia informatica si sta muovendo verso applicazioni online basate sul *cloud*, gli ostacoli associati all'accessibilità alle applicazioni per l'apprendimento della pronuncia scompariranno probabilmente in tempi brevi.

## Riferimenti bibliografici

- Añorga, A., & Benander, R. (2015). Creating a pronunciation profile of first-year Spanish students. *Foreign Language Annals*, 48(3), 434–446. <https://doi.org/10.1111/flan.12151>
- Aoyama, K., Guion, S., Flege, J. E., Yamada, T., & Akahane-Yamada, R. (2008). The first years in an L2-speaking environment: a comparison of Japanese children and adults learning American English. *International Review in Applied Linguistics*, 46(1), 61–90. <https://doi.org/10.1515/IRAL.2008.003>
- Audacity Team (2021). Audacity(R): Free Audio Editor and Recorder [Computer application]. Version 3.0.0 retrieved March 17th, 2021, from <https://audacityteam.org/>
- Baker, A., & Murphy, J. (2011). Knowledge base of pronunciation teaching: staking out the territory. *TESL Canada Journal*, 28(2), 29–50. <https://doi.org/10.18806/tesl.v28i2.1071>
- Baker, A. (2011). Discourse prosody and teachers' stated beliefs and practices. *TESOL Journal*, 2(3), 263–292. <https://doi.org/10.5054/tj.2011.259955>
- Barcomb, M., & Cardoso, W. (2020). Rock or lock? Gamifying an online course management system for pronunciation instruction: focus on English/r/and/l/. *CALICO Journal*, 37(2), 127–147. <https://doi.org/10.1558/cj.36996>
- Bradlow, A. R., Pisoni, D. B., Akahane-Yamada, R., & Tohkura, Y. (1997). Training Japanese listeners to identify English/r/and/l/: IV. Some effects of perceptual learning on speech production. *Journal of the Acoustic Society of America*, 101(4), 2299–2310. <https://doi.org/10.1121/1.418276>
- Boersma, P., & Weenink, D. (2022). Praat: doing phonetics by computer [Computer program]. Version 6.2.14, retrieved 24 May 2022 from <http://www.praat.org/>
- Carlet, A. F. (2017). L2 perception and production of English consonants and vowels by Catalan speakers: The effects of attention and training task in a cross-training study (Doctoral dissertation).
- Chun, D. M. (2013). Computer-assisted pronunciation teaching. In *Encyclopedia of applied linguistics*. C. A. Chapelle (Ed.), Oxford: Wiley-Blackwell, 823–834.
- Chun, D. M., Jiang, Y., Meyr, J., & Yang, R. (2015). Acquisition of L2 Mandarin Chinese tones with learner-created tone visualizations. *Journal of Second Language Pronunciation*, 1(1), 86–114. <https://doi.org/10.1075/jslp.1.1.04chu>
- Counselman, D. (2015). Directing attention to pronunciation in the second language classroom. *Hispania*, 98(1), 31–46. <https://doi.org/10.1353/hpn.2015.0006>
- Couper, G. (2017). Teacher cognition of pronunciation teaching: teachers' concerns and issues. *Tesol Quarterly*, 51(4), 820–843. <https://doi.org/10.1002/tesq.354>
- Derwing, T. M., Munro, M. J., & Wiebe, G. E. (1998). Evidence in favor of a broad framework for pronunciation instruction. *Language Learning*, 48(3), 393–410. <https://doi.org/10.1111/0023-8333.00047>
- Derwing, T. M., & Munro, M. J. (2009). Putting accent in its place: Rethinking obstacles to communication. *Language Teaching*, 42(4), 476–490. <https://doi.org/10.1017/S026-144480800551X>
- Derwing, T. M., Thomson, R. I., Foote, J. A., & Munro, M. J. (2012). A longitudinal study of listening perception in adult learners of English: implications for teachers. *Canadian Modern Language Review*, 68(3), 247–266. <https://doi.org/10.3138/cmlr.1215>
- Derwing, T. M., Munro, M. J., Foote, J. A., Waugh, E., & Fleming, J. (2014). Opening the

- window on comprehensible pronunciation after 19 years: a workplace training study. *Language Learning*, 64(3), 526–548. <https://doi.org/10.1111/lang.12053>
- Derwing, T. M., & Munro, M. J. (2015). *Pronunciation fundamentals: evidence-based perspectives for L2 teaching and research*. Amsterdam, Netherlands: John Benjamins.
- Engwall, O., Wik, P., Beskow, J., & Granström, B. (2004). Design strategies for a virtual language tutor. 8th international conference on spoken, Jiju Island, South Korea, March 31, 2004 (Seoul, South Korea: ISCA), 1–4.
- Evers, K., Chen, S. (2020). Effects of automatic speech recognition software on pronunciation for adults with different learning styles. *Journal of Educational Computing Research*, 59(4), 669–685. <https://doi.org/10.1177/0735633120972011>
- Foot, J., & McDonough, K. (2017). Using shadowing with mobile technology to improve L2 pronunciation. *Journal of Second Language Pronunciation*, 3(1), 34–56. <https://doi.org/10.1075/jslp.3.1.02foo>
- Haslam, M. (2011). The effect of perceptual training including required lexical access and meaningful linguistic context on second language phonology. PhD dissertation. Salt Lake City, Utah: University of Utah.
- Hardison, D. M. (2003). Acquisition of second-language speech: effects of visual cues, context, and talker variability. *Applied Psycholinguistics*, 24(4), 495–522. <https://doi.org/10.1017/S0142716403000250>
- Hardison, D. M. (2004). Generalization of computer-assisted prosody training: quantitative and qualitative findings. *Language Learning Technology*, 8(1), 34–52. <http://dx.doi.org/10.125/25228>
- Hazan, V., Sennema, A., Iba, M., & Faulkner, A. (2005). Effect of audiovisual perceptual training on the perception and production of consonants by Japanese learners of English. *Speech Communication*, 47(3), 360–378. <https://doi.org/10.1016/j.specom.2005.04.007>
- Hincks, R. (2003). Speech technologies for pronunciation feedback and evaluation. *ReCALL*, 15(1), 3–20. <https://doi.org/10.1017/S0958344003000211>
- Huensch, A., & Tremblay, A. (2015). Effects of perceptual phonetic training on the perception and production of second language syllable structure. *Journal of Phonetics*, 52, 105–120. <https://doi.org/10.1016/j.wocn.2015.06.007>
- Iverson, P., Pinet, M., & Evans, B. G. (2011). Auditory training for experienced and inexperienced second language learners: native French speakers learning English vowels. *Applied Psycholinguistics*, 33(1), 145–160. <https://doi.org/10.1017/S0142716411000300>
- Jenkins, J. (2004). Research in teaching pronunciation and intonation. *Annual Review of Applied Linguistics*, 24, 109–125. <https://doi.org/10.1017/S0267190504000054>
- Kaiser, D. J. (2017, June 8). IPronounce: Understanding pronunciation apps [Webinar]. Laureate-Cambridge Online Language Learning Research Network (OLLReN). Retrieved September 30th, 2022, from <https://ollren.org/%20events/past-events-s-lg-box-wrapper-17792711>
- Kennedy, S., & Trofimovich, P. (2010). Language awareness and second language pronunciation: a classroom study. *Language Awareness*, 19(3), 171–185. <https://doi.org/10.1080/09658416.2010.486439>
- Kholis, A. (2021). Elsa Speak App: Automatic Speech Recognition (ASR) for Supplementing English Pronunciation Skills. *Pedagogy: Journal of English Language Teaching*, 9(1), 1–14. <https://doi.org/10.32332/joelt.v9i1.2723>
- Lee, J., Jang, J., & Plonsky, L. (2014). The effectiveness of second language pronunciation instruction: a meta-analysis. *Applied Linguistics*, 36(3), 1–23. <https://doi.org/10.1093/applin/amu040>
- Lengeris, A. (2018). Computer-based auditory training improves second-language vowel production in spontaneous speech. *Journal of Acoustic Society of America*, 144(3). <https://doi.org/10.1121/1.5052201>
- Levis, J. (1999). Intonation in theory and practice, revisited. *Tesol Quarterly*, 33(1), 37–63. <https://doi.org/10.2307/3588190>
- Levis, J. (2007). Computer technology in teaching and researching. *Annual Review of Applied Linguistics*, 27, 184–202. <https://doi.org/10.1017/S0267190508070098>

- Levis, J. M. (2016). Research into practice: how research appears in pronunciation teaching materials. *Language Teaching*, 49(3), 423–437. <https://doi.org/10.1017/S0261444816000045>
- Levis, J. (2020). Revisiting the intelligibility and nativeness principles. *Journal of Second Language Pronunciation*, 6(3), 310–328. <https://doi.org/10.1075/jslp.20050.lev>
- Levis, J., & Pickering, L. (2004). Teaching intonation in discourse using speech visualization technology. *System*, 32, 505–524. <https://doi.org/10.1016/j.system.2004.09.009>
- Levis, J. M., Sonsaat, S., Link, S., & Barriuso, T. A. (2016). Native and nonnative teachers of L2 pronunciation: effects on learner performance. *Tesol Quarterly*, 50(4), 894–951. <https://doi.org/10.1002/tesq.272>
- Liakin, D., Cardoso, W., & Liakina, N. (2017). The pedagogical use of mobile speech synthesis (TTS): focus on French liaison. *Computer Assisted Language Learning*, 30(3–4), 348–365. <https://doi.org/10.1080/09588221.2017.1312463>
- Lima, E. F. (2015). Feel the rhythm! Fun and effective pronunciation practice using Audacity and sitcom scenes (teaching tip). In J. Levis, R. Mohammed, M. Qian, & Z. Zhou (Eds.), *Proceedings of the 6th pronunciation in second language learning and teaching conference, Santa Barbara* (pp. 277–284). Ames, IA: Iowa State University.
- Liu, X., Xu, M., Li, M. et al. (2019). Improving English pronunciation via automatic speech recognition technology. *International Journal of Innovation and Learning*, 25(2), 126–140. <https://doi.org/10.1504/IJIL.2019.097674>
- Martin, I. A. (2018). Bridging the gap between L2 pronunciation research and teaching: using iCPRs to improve German learners' pronunciation in distance and face-to-face classrooms [PhD dissertation]. State College, PA: The Pennsylvania State University
- McCrocklin, S.M. (2016). Pronunciation learner autonomy: The potential of Automatic Speech Recognition. *System*, 57, 25–42. <https://doi.org/10.1016/j.system.2015.12.013>
- McCrocklin, S., & Edalatshams, I. (2020). Revisiting popular speech recognition software for ESL speech. *Tesol Quarterly*, 54(4), 1086–1097. <https://doi.org/10.1002/tesq.3006>
- Melnik, G.A., & Peperkamp, S. (2021). High-Variability Phonetic Training enhances second language lexical processing: evidence from online training of French learners of English. *Bilingualism: Language and Cognition*, 24, 497–506. <https://doi.org/10.1017/S1366728920000644>
- Mitra, S., Tooley, J., Inamdar, P., & Dixon, P. (2003). Improving English pronunciation: An automated instructional approach. *Information Technologies and International Development*, 1(1), 75–84. <https://doi.org/10.1162/154475203771799720>
- Motohashi-Saigo, M., & Hardison, D. M. (2009). Acquisition of L2 Japanese geminates: training with waveform displays. *Language Learning Technology*, 13(2), 29–47. <https://doi.org/10.125/44179>
- Munro, M. J., & Derwing, T. M. (2006). The functional load principle in ESL pronunciation instruction: an exploratory study. *System*, 34(4), 520–531. <https://doi.org/10.1016/j.system.2006.09.004>
- Nagle, C. (2018). Motivation, comprehensibility, and accentedness in L2 Spanish: investigating motivation as a time varying predictor of pronunciation development. *Modern Language Journal*, 102(1), 199–217. <https://doi.org/10.1111/modl.12461>
- Neri, A., Mich, O., Gerosa, M., & Giuliani, D. (2008). The effectiveness of computer assisted pronunciation training for foreign language learning by children. *Computer Assisted Language Learning*, 21(5), 393–408. <https://doi.org/10.1080/09588220802447651>
- O'Brien, M. G. (2019). Targeting pronunciation (and perception) with technology. In N. Arnold, & L. Ducate (Eds.), *Engaging language learners through CALL*. Sheffield: Equinox, 309–352.
- O'Brien, M. G., Derwing, T. M., Cucchiari, C., Hardison, D. M., Mixdorff, H., Thomson, R., Strik, H., Levis, J. M., Munro, M. J., Foote, J. A., & Muller Levis, G. (2018). Directions for the future of technology in pronunciation research and teaching. *Journal of Second Language Pronunciation*, 4(2), 182–207. <https://doi.org/10.1075/jslp.17001.obr>
- Offerman, H. M., & Olson, D. J. (2016). Visual feedback and second language segmental production: the generalizability of pronunciation gains. *System*, 59, 45–60. <https://doi.org/10.1016/j.system.2016.03.003>
- Okuno, T., & Hardison, D. M. (2016). Perception–production link in L2 Japanese vowel du-



- ration: training with technology. *Language Learning Technology*, 20(2), 61–80. <https://doi.org/10.125/44461>
- Olson, D. J. (2019). Feature acquisition in second language phonetic development: evidence from phonetic training. *Language Learning*, 69(2), 366–404. <https://doi.org/10.1111/lang.12336>
- Olson, D. J., & Offerman, H. M. (2020). Maximizing the effect of visual feedback for pronunciation instruction: a comparative analysis of three approaches. *Journal of Second Language Pronunciation*, 7(1), 89–115. <https://doi.org/10.1075/jslp.20005.ols>
- Sakai, M., & Moorman, C. (2018). Can perception training improve the production of second-language phonemes? A meta-analytic review of 25 years of perception training research. *Applied Psycholinguistics*, 39(1), 187–224. <https://doi.org/10.1017/S0142-716417000418>
- Saito, K. (2013). Reexamining effects of form-focused instruction on L2 pronunciation development. *Studies in Second Language Acquisition*, 35(1), 1–29. <https://doi.org/10.1017/S0272263112000666>
- Saito, K., & Lyster, R. (2012). Effects of form focused instruction and corrective feedback on L2 pronunciation development of /f/ by Japanese learners of English. *Language Learning*, 62(2), 595–633. <https://doi.org/10.1111/j.1467-9922.2011.00639.x>
- Scales, J., Wennerstrom, A., Richard, D., & Wu, S. H. (2006). Language learners' perceptions of accent. *Tesol Quarterly*, 40, 715–738. <https://doi.org/10.2307/40264305>
- Sifakis, N. C., & Sougari, A. M. (2005). Pronunciation issues and EIL pedagogy in the periphery: a survey of Greek state schoolteachers' beliefs. *Tesol Quarterly*, 39, 467–488. <https://doi.org/10.2307/3588490>
- Suemitsu, A., Dang, J., Ito, T., & Tiede, M. (2015). A real-time articulatory visual feedback approach with target presentation for second language pronunciation learning. *Journal of the Acoustic Society of America*, 138(4). <https://doi.org/10.1121/1.4931827>
- Tejedor-García, C., Escudero-Mancebo, D., Cardeñoso-Payo, V., & González-Ferreras, C. (2020). Using challenges to enhance a learning game for pronunciation training of English as a second language. *IEEE Access*, 8, 74250–74266. <https://doi.org/10.1109/ACCESS.2020.2988406>
- Thomson, R. I. (2011). Computer-assisted pronunciation training: targeting second language vowel perception improves pronunciation. *CALICO Journal*, 28(3), 744–765. <https://doi.org/10.11139/cj.28.3.744-765>
- Thomson, R. I. (2018). High Variability [Pronunciation] Training (HVPT). A proven technique about which every language teacher and learner ought to know. *Journal of Second Language Pronunciation*, 4(2), 208–231. <https://doi.org/10.1075/jslp.17038.tho>
- Thomson, R. I., & Derwing, T. M. (2015). The effectiveness of L2 pronunciation instruction: a narrative review. *Applied Linguistics*, 36(3), 326–344. <https://doi.org/10.1093/applin/amu076>
- Trofimovich, P., Iacobs, T., Kennedy, S., Saito, K., & Crowther, D. (2016). Flawed self-assessment: investigating self-and other perception of second language speech. *Bilingualism: Language and Cognition*, 19(1), 122–140. <https://doi.org/10.1017/S1366728914000832>
- Wallace L. (2015). Using Google web speech as a springboard for identifying personal pronunciation problems. The 7th Annual Pronunciation in Second Language Learning and Teaching Conference, Dallas, USA.
- Wang, X. (2002). *Training Mandarin and Cantonese speakers to identify English vowel contrasts: long-term retention and effects on production*. Burnaby, Canada: Simon Fraser University.
- Yuan, Y., Liu, X. (2020). An empirical study of the effect of ASR-supported English reading aloud practices on pronunciation accuracy. In The 5th International Conference on Technology in Education (ICTE 2020), Macau, China.
- Zajc, M., & Rojczyk, C. (2014). Imitation of English vowel duration upon exposure to native and non-native speech. *Poznań Studies in Contemporary Linguistics*, 50(4), 495–514. <https://doi.org/10.1515/psicl-2014-0025>