

FEDERICA CAVICCHIO, MIRKO GRIMALDI

## The Effect of Subtitles on Second Language Pronunciation

The increasing prevalence of video-on-demand services for watching movies and TV shows presents a unique opportunity for foreign language learning, thanks to easy access to original language content. When watching films in the original language, viewers typically can follow the audiovisual content with subtitles either in their language (L1) or the film's language (L2). Here, we present a pilot study that examines the impact of English (L2) subtitles, Italian (L1) subtitles, or no subtitles on L1 native Italian speakers' production of L2 English phonemes. Our findings reveal that subtitles in the L2 language (English) can improve the production of L2 phonemes. However, while L2 subtitles facilitated the improved pronunciation of the vowel /æ/, no significant enhancements were observed for /ʌ/. The study emphasizes the utility of incorporating reading and listening in L2 for effective language acquisition. It also points out the importance of considering L1 phonemic traits that may pose challenges to L2 acquisition.

*Keywords:* Audiovisual online content, Subtitles, L2 acquisition, L1-L2 interference.

### 1. *Introduction*

The critical period hypothesis (CPH, see, for example, Lennenberg, 1967; Pinker, 1984; Hartshorne, Tenenbaum & Pinker, 2018; MacWhinney, 1987) posits that there is a specific window in human development during which acquisition of the native language (L1) occurs most naturally and efficiently. This period is typically considered to be during early childhood. After this window closes, it is believed that the ability to learn a new language (L2) with native-like fluency significantly decreases. This notion is often juxtaposed with neuroplasticity (see, for example, Merzenich, Nahum, & Van Vleet, 2013), that is, the capacity of the brain to form new neural connections throughout an individual's life in response to new experiences and learning.

During the critical period, children's brains exhibit heightened neuroplasticity (e.g., Kuhl, Conboy, Padden, Nelson & Pruitt, 2005; Kuhl, Conboy, Coffey-Corina, Padden, Rivera-Gaxiola & Nelson, 2008; Morford, Mayberry, 2000), allowing for more efficient development of neural pathways responsible for language processing and production. This heightened plasticity enables children to assimilate new phonological systems more readily, particularly when exposed to an L2 through native speakers, even in the context of the classroom (see Cheour, Shestakova, Alku, Ceponiene & Näätänen, 2002; Shestakova, Huotilainen, Čeponien & Cheour, 2003; Peltola, Kuntola, Tamminen, Hämäläinen & Aaltonen, 2005).

Neuroplasticity allows for continuous learning and adaptation within the human brain, including acquiring an L2 at different ages (Stein, Kuntola, Tamminen, Hämäläinen & Aaltonen, 2006; Leagault, Grant, Fang & Li, 2019a; Leagault, Fang, Lan & Li, 2019b). Neuroplasticity underpins the processes through which adult learners can immerse themselves in an L2 environment and experience a range of linguistic inputs that challenge them to categorise and produce new L2 sounds, as described by Flege (1995), Flege, Bohn (2021) and Winkler, Kujala, Tiitinen, Sivonen, Alku, Lehtokoski, Czigler, Csépe, Ilmoniemi & Näätänen (1999). The ability to leverage neuroplasticity for L2 learning, even in adults, underlines the brain's capacity to reorganise and adjust to new linguistic contexts, even outside the critical period. Several factors might amplify L2 learning, including linguistic aptitude (Hu, Ackermann, Martin, Erb, Winkler & Reiterer, 2013; Chai, Berken, Barbeau, Soles, Callahan & Chen, 2016), along with the intensity (Thomson, Derwing, 2015) and quality (Zhang, Kuhl, Imada, Iverson, Pruitt, Stevens & Nemoto, 2009) of L2 learning. Conversely, when L2 acquisition occurs in a classroom environment where the teachers have a non-native accent, and the primary language of communication among students remains their L1, the immersion necessary for this neural activation is insufficient. In such environments, the exposure to native L2 phonetics and syntax is significantly reduced, and students may inadvertently adopt their non-native teachers' inaccurate pronunciations and rhythms. Furthermore, the prevalence of L1 during classroom activities may lead to a reliance on familiar linguistic structures, preventing the necessary neural adaptation and integration of L2 linguistic patterns (Grimaldi, Sisinni, Gili Fivela, Invitto, Resta & Alku, 2014).

Despite this inherent potential for neuroplastic adaptation to new linguistic environments, longitudinal studies on auditory plasticity have shown the challenges involved in L2 phonetic learning within the constraints of traditional classroom settings. These studies, such as those conducted by Grimaldi et al. (2014), Jost, Eberhard-Moscicka, Pleisch, Heusser, Brandeis, Zevin & Maurer (2015), Hisagi, Shafer, Miyagawa, Kotek, Sugawara & Pantazis (2016), Wottawa, Adda-Decker & Isel (2022), and Højlund, Horn, Sørensen, McGregor & Wallentin (2022), indicate that sustained exposure to an L2 in a classroom environment does not necessarily result in changes to auditory plasticity. These results suggest that, over time, the auditory system's sensitivity to the acoustic nuances of a new language may not be significantly enhanced in such settings. As posited by Piske (2007), one potential explanation for this finding is that the quality and intensity of L2 stimuli provided in classrooms are insufficient to reinvigorate the auditory system's sensitivity to the spectro-temporal characteristics unique to non-native languages. In a typical classroom setting, the auditory experience of L2 may be diluted by various factors: the presence of non-native accents, limited opportunity for immersive interaction, and the predominance of controlled, repetitive exercises over spontaneous language use. These findings have considerable implications for the formation of mental representations of L2 phonemes. The ability to create accurate mental models of

L2 phonemic structures is vital for successful L2 acquisition, but when the auditory input lacks richness and authenticity or when exposure to the language is fragmented or inconsistent, the establishment of clear and distinct phonemic representations may be impeded.

Classroom learning may challenge the development of L2 representations for several reasons. First, the frequency of exposure to L2 sounds may be limited to a few hours per week, which is insufficient compared to the constant stimulation provided in a naturalistic environment. Second, the diversity of L2 input in a classroom is often narrower than in real-world settings, offering fewer opportunities for learners to experience the full range of phonetic variations and contextual uses of the language. In summary, the dynamics of a classroom may not always assist the perceptual and learning processes essential for refining the perception and production of L2 phonemes. Regarding L2 sound acquisition, recent meta-analytic reviews (e.g., Sakai, Moorman, 2018; Rato, Oliveira, 2023) have robustly validated that the use of brief training sessions focused on L2 sound perception (e.g. High Variability Phonetic Training, HVPT) can enhance the identification and production of L2 phonemes. In light of these findings, we must rethink L2 teaching methodologies incorporating methods that can enhance auditory plasticity, such as immersive experiences and short acoustic training, and technologies that can simulate natural language exposure.

The need for innovative immersive learning strategies resonates with the growing use of video-on-demand services, which provide access to content in its original language accompanied by subtitles in either the learners' native language (L1) or the original language of the film (L2). In a seminal study, Mitterer, McQueen (2009) explored the effects of either L1 or L2 subtitles on Dutch L2 speakers' perception of Scots English words after watching the film "Trainspotting". Their findings suggested that exposure to L2 subtitles could enhance speech comprehension more than L1 subtitles. Subsequent studies have confirmed that using L2 subtitles significantly facilitates L2 acquisition. A study by Vulchanova, Aurstad, Kvitnes & Eshuis (2015) demonstrates that L2 subtitles are more effective than L1 subtitles in aiding the acquisition of L2 vocabulary. Frumuselu, De Maeyer, Donche & Colon Plana (2015) found that L2 subtitles contribute to a better long-term understanding of spoken L2. These results are likely due to the direct visual reinforcement of spoken words and the active engagement of both auditory and visual processing pathways, which enriches the language learning experience and aids the processing of language constructs. The simultaneous exposure to spoken and written forms of the target language enables learners to make connections between the phonetic sounds of words and their orthographic representations, fostering a more comprehensive grasp of the language as used in natural contexts (on the subject of L2 phono-lexical representation and phoneme/grapheme interference see Darcy, Holliday, 2019; and Bassetti, Cerni & Masterson, 2022, respectively).

Expanding on the effect on L2 perception of L2 subtitles, Wisniewska, Mora (2021) report that exposure to L2 subtitles not only can enhance L2 phoneme

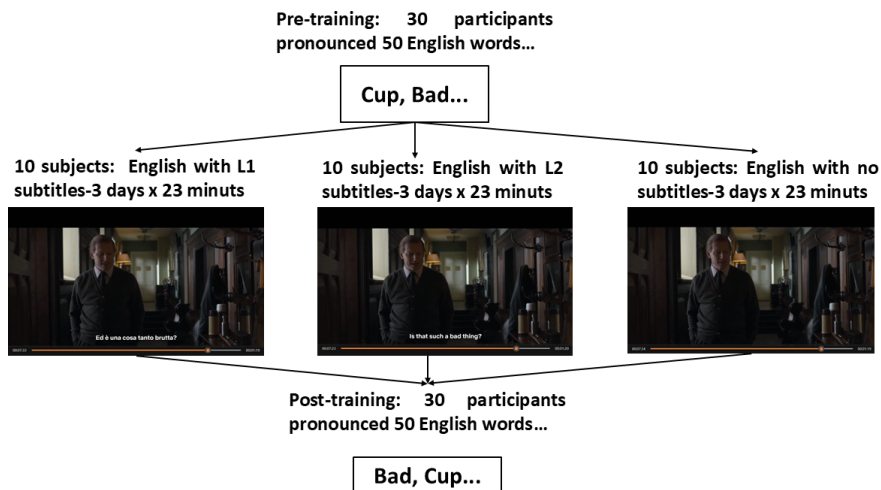
perception, effectively helping L2 learners to distinguish between similar words, but can also improve the learners' accent, as per native speakers' ratings of L2 speech productions. However, it is crucial to note that while subtitles may improve the perception of L2 pronunciation, they do not necessarily translate into the learners' ability to produce all the sounds accurately, as reported by Hutchinson, Dmitrieva (2022), who point out that exposure to L2 subtitles does not improve L2 vowel production, as per native speakers' ratings.

In summary, according to results from previous studies, L2 subtitles, while widely acknowledged for their efficacy in enhancing language comprehension and vocabulary acquisition, may not directly facilitate the accurate production of L2 phonemes. Our pilot study aims to test whether exposure to L2 subtitles can influence the pronunciation skills of native Salento Italian speakers learning English as L2. Our study specifically investigates the impact of different subtitle settings on the acquisition of British English phonemes that are commonly mispronounced by Italian speakers from Salento. We selected the phonemes /ʌ/ (as in "cup") and /æ/ (as in "bad"), which are often confused by Salento Italian speakers with the Italian vowels /a/ and /ɛ/, respectively (Escudero, Sisinni, & Grimaldi, 2014). English subtitles might provide auditory and visual reinforcement of the phonetic characteristics of English, potentially aiding learners in overcoming the influence of their L1 phonemes. By contrast, Italian subtitles might reinforce reliance on native pronunciation norms, whereas the no subtitles condition, should not help or hinder L2 pronunciation. Our method included pre- and post-tests to evaluate the accuracy of phoneme production before and after participants were exposed to one of the three subtitles conditions. Their L2 productions were compared to the productions of native English speakers. We were able to make this direct comparison because, unlike previous studies, we do not employ qualitative assessments by L1 speakers to assess the participants' L2 productions. Instead, to quantify the deviation of L2 speakers' vowel pronunciation from that of native English speakers, we utilised the Mahalanobis distance (Mahalanobis, 1936), a statistical measure useful in the context of measuring how similar or different a given observation is from a reference group, especially when the data is multidimensional. Unlike other metrics, such as the Euclidean distance, the Mahalanobis distance considers the correlations across multidimensional data. Therefore, the Mahalanobis distance is particularly suitable for analysing vowel existence spaces. Vowels are characterised by their formant frequencies, specifically the first and second formants (F1 and F2), which are multidimensional and interrelated. The Mahalanobis distance provides a robust method for measuring distances within this complex space. Consequently, it enabled us to precisely gauge how closely or not the vowel pronunciations of L2 speakers matched those of native English speakers, considering variations within each subtitles group. Unlike previous studies, our approach provides a quantitative, rather than qualitative, assessment of the L2 speakers' phonetic variation.

## 2. Materials and Methods

Thirty native Italian speakers from the Salento region (mean age 23.4 years, 23 women) participated in the study. They were all first-year students at the Faculty of Foreign Languages and Literatures, and their English level, according to CEFR, was B1. It should be noted that a five-vowel system characterises the local dialect in Salento (/i, e, a, ɔ, u/), with noted variations as detailed by Grimaldi (2003) and Romano (2013). We selected two segments from the Netflix series “The Crown,” specifically from Season 1, Episode 2 and Season 4, Episode 8, which together amounted to approximately 23 minutes of viewing. From these episodes, we selected twenty-five words containing the vowel /ʌ/ and twenty-five words containing the vowel /æ/ in stressed positions. The words were chosen for their similar frequency of usage, as reported in the *BNC XML*. Five female actresses and five male actors pronounced the words in the film segments, and the English variant spoken in “The Crown” is the Received Pronunciation (RP). In Fig. 1, we present a sketch of the experimental procedure.

Figure 1 - *Experimental Procedure Overview. The experimental protocol began with the pre-test. Participants recorded 50 words, 25 words containing the vowel /æ/ (as in “bad”) and 25 words with the vowel /ʌ/ (as in “cup”). The words were selected from the film segments used later in the study. Participants were then assigned to one of three subtitle conditions and then re-tested in the post-test*



Participants recorded the words selected from the TV series segments in a soundproof booth. The words were displayed one at a time in written format via MS PowerPoint on a Lenovo portable PC. Participants articulated the words at their own pace. The recording sessions lasted 10 minutes on average and were conducted using Audacity software version 3.08, with a Trust 21674 Table Microphone. Subsequently, participants were divided into three groups: the first

group watched the film segments with L2 English subtitles, the second with L1 Italian subtitles, and the third without subtitles. Participants viewed the segments once daily for three consecutive days via the online experiment platform *Gorilla*. Following this exposure, participants returned to record the same words presented in the pre-test, which were presented to them in a counterbalanced order with respect to the initial recording.

### 3. *Data Analysis*

We extracted the F1 and F2 formant values of the / $\Delta$ / and / $\text{æ}$ / vowels articulated by the 30 L2 participants and 10 L1 speakers using DARTmouth Linguistic Automation (DARLA) software (Reddy, Stanford, 2015). To assess the vowels' production differences between L2 and L1 speakers, we computed the Mahalanobis distance (Mahalanobis, 1936) using package *dplyr* (Wickham, François, Henry, Müller & Vaughan, 2023) in R (version 4.2). We calculated the mean vectors and the covariance matrix for F1 and F2 values for each vowel and each speaker and then computed the Mahalanobis distance between L2 and L1 speakers. Smaller Mahalanobis distance values indicate a higher similarity between the vowel spaces of L2 speakers and L1 speakers, suggesting a closer approximation to native-like vowel production by the L2 speakers. Conversely, larger Mahalanobis distances point to significant discrepancies in vowel production between L2 and L1 speakers. By comparing the Mahalanobis distance values of the pre-test and post-test under different subtitle conditions, we can gauge whether a phonological adjustment may have occurred for L2 speakers due to exposure to different subtitle conditions.

### 4. *Results and Discussion*

Tab. 1 reports the statistical summary of Mahalanobis distances for the vowels / $\Delta$ / and / $\text{æ}$ / measured before (pre-test) and after (post-test) the participants were exposed to three different subtitle conditions. The summary includes the mean, maximum, minimum, and median Mahalanobis distance values. The results indicate a reduction in Mahalanobis distances for the vowel / $\text{æ}$ / following exposure to subtitles. Specifically, the mean Mahalanobis distance for / $\text{æ}$ / in the pre-test was 2.23; this distance decreased to 1.65 after exposure to L1 subtitles and an even lower mean value of 0.91 when subtitles were in the L2. In contrast, the condition with no subtitles showed no substantial change in the mean value of the Mahalanobis distance (mean=2.6). Conversely, the phoneme / $\Delta$ / exhibited no significant variation in Mahalanobis distances across different subtitle conditions, indicating that the efficacy of subtitle-based phonetic training may be vowel-specific. This finding suggests that the impact of subtitles on phonetic learning is influenced not only by the language of the subtitles but also by the phonemes involved, with L2 subtitles showing greater effectiveness for the vowel / $\text{æ}$ / but not for / $\Delta$ /.

Table 1 - Summary of Mahalanobis Distance Measures for Vowels /ʌ/ and /æ/ comparing values from the pre-test and post-test across different subtitle conditions. The mean Mahalanobis distance values lower in the post-test, indicating greater proximity to native speakers' vowels, are highlighted in bold print

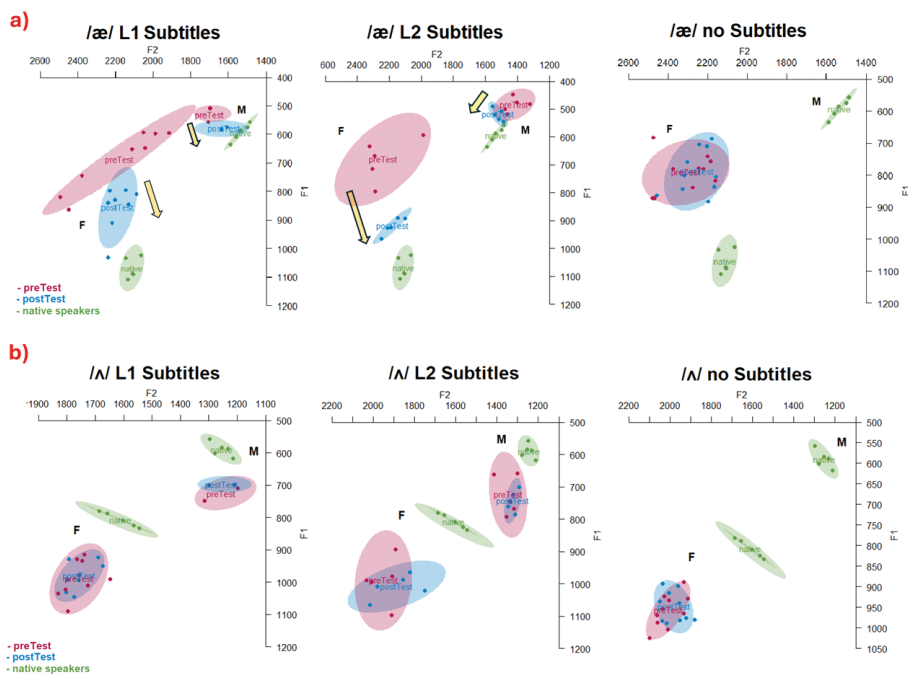
|     |        | Pre-test | Post-test<br>L2 Subtitles | Post-test<br>L1 Subtitles | Post-test<br>No Subtitles |
|-----|--------|----------|---------------------------|---------------------------|---------------------------|
| /ʌ/ | max    | 9.48     | 6.93                      | 6.6                       | 11.8                      |
|     | min    | 0.02     | 0.02                      | 0.02                      | 0.07                      |
|     | mean   | 2.25     | 2.04                      | 2.06                      | 2.28                      |
|     | median | 1.97     | 1.6                       | 2                         | 2.32                      |
| /æ/ | max    | 11.65    | 5.92                      | 9.3                       | 14.72                     |
|     | min    | 0.02     | 0.004                     | 0.009                     | 0.08                      |
|     | mean   | 2.23     | 0.91                      | 1.65                      | 2.6                       |
|     | median | 1.58     | 0.65                      | 1.01                      | 1.82                      |

In Fig. 2, we present a visual assessment of participants' different subtitle conditions for vowel production. The figure presents the existence areas (F1x F2) for vowels /æ/ and /ʌ/, comparing data collected during the pre-test and post-test sessions. The analysis is differentiated by subtitle conditions (L1 subtitles, L2 subtitles, and no subtitles). Additionally, the figure delineates the vowel existence areas for both female and male speakers. Participants consistently undershoot (Flege, Schirru & MacKay, 2003) across conditions, that is, participants did not fully achieve the articulatory targets of the English vowels and consistently fell short of the precise phonetic qualities of the target English sounds. However, for the vowel /æ/, speakers in the L2 subtitles condition showed a notable shift in their vowel existence areas towards those typical of native speakers.

The findings of this pilot study show that the subtitles' role in L2 production is nuanced. Our results revealed that exposure to L2 subtitles improved the vowel /æ/ production, as evidenced by the reduced Mahalanobis distances in the post-test data for the L2 subtitles' group. This improvement aligns with previous observation that English subtitles might facilitate the L2 learners' ability to differentiate and accurately produce non-native phonemes. In particular, our results are congruent with the findings of Wisniewska, Mora (2021), which suggested that L2 subtitles can enhance phoneme perception and even improve accent in L2 learners. Using the Mahalanobis distance, our study adds a quantitative dimension to Wisniewska and Mora (2021) qualitative observations, demonstrating that L2 subtitles' impact on L2 speech production is selective and varies by phoneme. As Hutchinson, Dmitrieva (2022) noted, exposure to L2 subtitles does not necessarily translate into an equivalent L2 speech production enhancement for all the L2 vowels. Indeed, in our study, the phoneme /ʌ/ showed no substantial changes in the Mahalanobis distances across any of the subtitle conditions, indicating a potential limitation in the impact of subtitles on the acquisition and production of that specific phonemes.

The “resistance to change” for /ʌ/ could depend on a stronger L1 interference for /ʌ/ and /æ/.

Figure 2 - Existence areas (F1XF2) for vowels /æ/ (panel a) and /ʌ/ (panel b) in the pre-test and post-test by subtitle conditions (L1 subtitles, L2 subtitles, and no subtitles). Vowel existence areas are pinpointed for female (F) and male (M) speakers



Previous studies by Cavicchio, Grimaldi (2022) have found that Salento Italian speakers consistently identify the vowel /ʌ/ less accurately than the vowel /ɒ/. In Salento Italian, a five-vowel system, lip rounding is not a distinctive feature, unlike frontness/backness and the height trait. Indeed, the trait [±rounded] in Salento Italian is redundant, as the vowels /ɔ, u/, compared to /i, ε/, are both [+back] and [+rounded], and the vowel /a/ is characterised solely by the [+low] trait: hence the traits [+low] and [-rounded] attributable to /a/ are predictable (Grimaldi, 2019). This difference could explain why acquiring /ʌ/ by speakers of Salento Italian is challenging. The lack of rounding as a distinctive phonological feature can impede the perception and production of /ʌ/, leading to a tendency to substitute /ʌ/ with the nearest L1 vowels, /a/.

Finally, the lack of significant improvement in vowel production under the Italian subtitles and the no subtitles conditions reinforces that L2 language exposure is more beneficial than reliance on L1 norms or the absence of linguistic cues. This finding is consistent with prior research by Mitterer, McQueen (2009). It also supports the theoretical framework posited by Vulchanova, Aurstad, Kvitnes

& Eshuis (2015), which emphasises the value of direct L2 input for its acquisition as opposed to instructions on L2 in the learners' native language.

Our study has some obvious limitations. First, it is a pilot that involves a relatively small cohort of participants, which can significantly constrain the generalizability of the findings. Due to the small number of observations, we did not include test statistics in our analysis. This choice reflects the nature of our study, which was intended to explore potential trends rather than provide definitive conclusions. Moreover, the predominance of female participants may introduce a bias, i.e., the results may not adequately represent how male learners, or a more balanced mix of female and male speakers, would respond to subtitle exposure. Future studies should aim to recruit a larger and more diverse participant pool to enhance the representativeness and reliability of the findings and allow for statistical analyses.

### 5. *Conclusions*

This pilot study contributes to the ongoing collection of evidence on the effectiveness of subtitles in L2 phonetic training by offering quantitative evidence of their possible benefits and underlying possible limitations regarding L2 speech production. Our findings contribute to the piling of evidence on the role of subtitle exposure in L2 learning. However, the selective efficacy of subtitles in improving the production of certain L2 vowels highlights the complex interplay between phonological features in L2 language learning. Ultimately, our results show the importance of considering the learners' native language phonological system when designing pedagogical strategies for L2 acquisition. Future studies should expand the cohort of L2 learners tested to provide a deeper understanding of the role of subtitles in L2 learning and explore the possibility of combining different forms of L2 training to address the acquisition of the more difficult phonemes.

### *References*

- BASSETTI, B., CERNI, T. & MASTERSON, J. (2022). The efficacy of grapheme-phoneme correspondence instruction in reducing the effect of orthographic forms on second language phonology. In *Applied Psycholinguistics*, 43(3), 683–705.
- BNC XML EDITION: <http://www.natcorp.ox.ac.uk/XMLedition/>
- CAVICCHIO, F., GRIMALDI, M. (2022). Enhancing L2 Speech Perception Using Auditory and Audiovisual Training. In AMLaP 28, University of York, 7-9 September 2022.
- CHAI, X.J., BERKEN, J.A., BARBEAU, E.B., SOLES, J., CALLAHAN, M. & CHEN, J.-K. (2016). Intrinsic functional connectivity in the adult brain and success in second-language learning. In *Journal of Neuroscience*, 36, 755–761.
- CHEOUR, M., SHESTAKOVA, A., ALKU, P., CEPONIENE, R. & NÄÄTÄNEN, R. (2002). Mismatch negativity shows that 3–6-year-old children can learn to discriminate non-native speech sounds within two months. In *Neuroscience Letters*, 325(3), 187–190.

- DARCY, I., HOLLIDAY, J.J. (2019). Teaching an old work new tricks: Phonological updates in the L2 mental lexicon. In LEVIS, J., NAGLE, C. & TODEY, E. (Eds.), *Proceedings of the 10th Pronunciation in Second Language Learning and Teaching Conference*, Ames, IA, September 2018, 10–26.
- ESCUADERO, P., SISINNI, B. & GRIMALDI, M. (2014). The effect of vowel inventory and acoustic properties in Salento Italian learners of Southern British English vowels. In *Journal of the Acoustic Society of America*, 135(3), 1577–1584.
- FLEGE, J.E. (1995). Second-language speech learning: Theory, findings, and problems. In STRANGE, W. (Ed.), *Speech perception and linguistic experience: Issue in cross-language research*. Timonium, MD: York, 229–273.
- FLEGE, J.E., SCHIRRU, C. & MACKAY, I.R. (2003). Interaction between the native and second language phonetic subsystems. In *Speech Communication*, 40(4), 467–491.
- FLEGE, J.E., BOHN, O-S. (2021). The Revised Speech Learning Model (SLM-r). In: WAYLAND, R. (Ed.). *Second Language Speech Learning: Theoretical and Empirical Progress*. Cambridge: Cambridge University Press, 3–83.
- FRUMUSELU, A.D., DE MAEYER, S., DONCHE, V. & GUTIERREZ COLON-PLANA, M. (2015). Television series inside the EFL classroom: Bridging the gap between teaching and learning informal language through subtitles. In *Linguistics and Education*, 32, 107–117.
- GORILLA EXPERIMENT SOFTWARE: [www.gorilla.sc](http://www.gorilla.sc)
- GRIMALDI, M., SISINNI, B., GILI FIVELA, B., INVITTO, S., RESTA, D. & ALKU, P. (2014). Assimilation of L2 vowels to L1 phonemes governs L2 learning in adulthood: a behavioral and ERP study. In *Frontiers in Human Neuroscience*, 8, 279.
- GRIMALDI, M. (2003). *Nuove ricerche sul vocalismo tonico del Salento meridionale. Analisi acustica e trattamento fonologico dei dati*. Alessandria: Edizioni dell'Orso.
- GRIMALDI, M. (2019). *Il cervello fonologico*. Roma: Carocci Editore.
- HARTSHORNE, J.K., TENENBAUM, J.B. & PINKER, S. (2018). A critical period for second language acquisition: Evidence from 2/3 million English speakers. In *Cognition*, 177, 263–277.
- HISAGI, M., SHAFER, V.L., MIYAGAWA, S., KOTEK, H., SUGAWARA, A. & PANTAZIS, D. (2016). Second-language learning effects on automaticity of speech processing of Japanese phonetic contrasts: An MEG study. In *Brain Research*, 1652, 111–118.
- HØJLUND, A., HORN, N.T., SØRENSEN, S.D., MCGREGOR, W.B. & WALLENTIN, M. (2022). Foreign language learning and the mismatch negativity (MMN): A longitudinal ERP study. In *Neuroimage: Reports*, 2(4), 100138.
- HU, X., ACKERMANN, H., MARTIN, J.A., ERB, M., WINKLER, S. & REITERER, S.M. (2013). Language aptitude for pronunciation in advanced second language (L2) Learners: Behavioural predictors and neural substrates. In *Brain and Language*, 127, 366–376.
- HUTCHINSON, A.E., DMITRIEVA, O. (2022). Exposure to speech via foreign film and its effects on non-native vowel production and perception. In *Journal of Phonetics*, 95, 101189.
- JOST, L.B., EBERHARD-MOSCICKA, A.K., PLEISCH, G., HEUSSER, V., BRANDEIS, D., ZEVIN, J.D. & MAURER, U. (2015). Native and non-native speech sound processing and the neural mismatch responses: A longitudinal study on classroom-based foreign language learning. In *Neuropsychologia*, 72, 94–104.

- KUHL, P.K., CONBOY, B.T., COFFEY-CORINA, S., PADDEN, D., RIVERA-GAXIOLA, M. & NELSON, T. (2008). Phonetic learning as a pathway to language: new data and native language magnet theory expanded. In *Philosophical Transactions of the Royal Society of London B Biological Science*, 363, 979–1000.
- KUHL, P.K., CONBOY, B.T., PADDEN, D., NELSON, T. & PRUITT, J. (2005). Early speech perception and later language development: implications for the “critical period”. In *Language Learning Development*, 1, 237–264.
- LEGAULT, J., GRANT, A., FANG, S.Y. & LI, P. (2019a). A longitudinal investigation of structural brain changes during second language learning. In *Brain and Language*, 197, 104661.
- LEGAULT, J., FANG, S.Y., LAN, Y.J. & LI, P. (2019b). Structural brain changes as a function of second language vocabulary training: Effects of learning context. In *Brain and Cognition*, 134, 90–102.
- LENNEBERG, E.H. (1967). *Biological foundations of language*. New York: Wiley.
- MAHALANOBIS, P.C. (1936). On the generalised distance in statistics. In *Proceedings of the National Institute of Sciences of India*, 2(1), 49–55.
- MACWHINNEY, B. (1987). Toward a psycholinguistically plausible parser. In THOMASON, S. (Ed.), *Proceedings of the Eastern States Conference on Linguistics*. Columbus, OH: Ohio State University, 12–32.
- MERZENICH, M.M., NAHUM, M. & VAN VLEET, T.M. (2013). Neuroplasticity: introduction. In *Progress in Brain Research*, 2(7), xxi–xxvi.
- MITTERER, H., MCQUEEN, J.M. (2009). Foreign Subtitles Help but Native-Language Subtitles Harm Foreign Speech Perception. In *PLoS ONE*, 4(11), e7785.
- MORFORD, J., MAYBERRY, R. (2000). *A reexamination of “early exposure” and its implications for language acquisition by eye*. Lawrence Erlbaum Associates.
- PINKER, S. (1984). *Language Learnability and Language Development*. Cambridge, MA: Harvard University Press.
- REDDY, S., STANFORD, J.N. (2015). Toward Completely Automated Vowel Extraction: Introducing DARLA. In *Linguistics Vanguard*, 1(1), 15–28.
- PELTOLA, M.S., KUNTOLA, M., TAMMINEN, H., HÄMÄLÄINEN, H. & AALTONEN, O. (2005). Early exposure to non-native language alters preattentive vowel discrimination. In *Neuroscience Letters*, 388(3), 121–125.
- PISKE, T. (2007). Implications of James E. Flege’s research for the foreign language classroom. In MUNRO, M.J., BOHN, O-S. (Eds.), *Language experience in second language speech learning. In honor of James Emil Flege*. Amsterdam: John Benjamins, 301–314.
- RATO, A., OLIVEIRA, D. (2023). Assessing the robustness of L2 perceptual training: A closer look at generalisation and retention of learning. Second Language Pronunciation. In ALVES, U.K., ALCANTARA DE ALBUQUERQUE, J.I. (Eds.), *Different Approaches to Teaching and Training*, Berlin, Boston: De Gruyter Mouton, 369–396.
- ROMANO, A. (2013). Il vocalismo del dialetto salentino di Galàtone: differenze di apertura metafonetiche, tracce isolate di romanzo comune o interferenze diasistematiche? In ROMANO, A., SPEDICATO, M. (Eds.), *Sub voce Sallentinitas. Studi in onore di p. Giovan Battista Mancarella*. Lecce: Edizioni del Grifo, 247–276.

- 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. In *Applied Psycholinguistics*, 39(1), 187–224.
- SHESTAKOVA, A., HUOTILAINEN, M., ČEPONIEN, R. & CHEOUR, M. (2003). Event-related potentials associated with second language learning in children. In *Clinical Neurophysiology*, 114(8), 1507–1512.
- STEIN, M., DIERKS, T., BRANDEIS, D., WIRTH, M., STRIK, W. & KOENIG, T. (2006). Plasticity in the adult language system: a longitudinal electrophysiological study on second language learning. In *Neuroimage*, 33(2), 774–783.
- THOMSON, R.I., DERWING, T.M. (2015). The Effectiveness of L2 Pronunciation Instruction: A Narrative Review. In *Applied Linguistics*, 36(3), 326–344.
- VULCHANOVA, M., AURSTAD, L.M., KVITNES, I.E. & ESHUIS, H. (2015). As naturalistic as it gets: Subtitles in the English classroom in Norway. In *Frontiers in Psychology*, 5, 98076.
- WICKHAM, H., FRANÇOIS, R., HENRY, L., MÜLLER, K. & VAUGHAN, D. (2023). *dplyr: A Grammar of Data Manipulation*. R package version 1.1.4, <https://github.com/tidyverse/dplyr>
- WINKLER, I., KUJALA, T., TIITINEN, H., SIVONEN, P., ALKU, P., LEHTOKOSKI, A., CZIGLER, I., CSÉPE, V., ILMONIEMI, R.J. & NÄÄTÄNEN, R. (1999). Brain responses reveal the learning of foreign language phonemes. In *Psychophysiology*, 36(5), 638–642.
- WISNIEWSKA, N., MORA, J.C. (2020). Can Captioned Video Benefit Second Language Pronunciation? In *Studies in Second Language Acquisition*, 42(3), 599–624.
- WOTTAWA, J., ADDA-DECKER, M. & ISEL, F. (2022). Neurophysiology of non-native sound discrimination: Evidence from German vowels and consonants in successive French–German bilinguals using an MMN oddball paradigm. In *Bilingualism: Language and Cognition*, 25, 137–147.
- ZHANG, Y., KUHL, P.K., IMADA, T., IVERSON, P., PRUITT, J., STEVENS, E.B. & NEMOTO, I. (2009). Neural signatures of phonetic learning in adulthood: A magnetoencephalography study. In *Neuroimage*, 46, 226–240.