

PLENARY SPEECH

Phonological acquisition in instructed SLA: L2 pronunciation training in and beyond the classroom and individual differences

Joan C. Mora

Department of Modern Languages and Literatures and English Studies. Faculty of Philology and Communication,
Universitat de Barcelona, Barcelona, Spain
Email: mora@ub.edu

(Received 21 July 2025; accepted 20 October 2025)

Abstract

This article provides an overview of key challenges in second language (L2) pronunciation learning and teaching within the context of instructed second language acquisition (SLA), with the goal of identifying promising directions for future research. It begins by examining persistent difficulties in L2 pronunciation instruction, such as the typically limited quality of input and the dominant emphasis on grammar and vocabulary in communicative language teaching (CLT). These conditions often result in learners having limited awareness of their pronunciation needs and teachers facing challenges in incorporating pronunciation instruction into CLT-based curricula. The article then reviews emerging instructional approaches that aim to integrate attention to phonetic form within CLT, highlighting the need for further empirical investigation. In addition, several pronunciation training techniques, some underexplored (HVPT, shadowing, embodied pronunciation training, captioned video, accent imitation, and pronunciation self-assessment), are briefly described, with an emphasis on their pedagogical potential both inside and outside the classroom. Finally, the article considers the role of individual differences in L2 pronunciation development and proposes directions for future research in instructed SLA.

Keywords: pronunciation instruction; pronunciation training techniques; phonetic training; task-based pronunciation teaching (TBPT); individual differences

1. Introduction: L2 pronunciation is a teaching and learning challenge

Learning and teaching foreign and second language (L2) pronunciation is a major challenge in instructed second language acquisition (SLA). For many learners, pronunciation is difficult, and often a learning challenge. Survey studies have shown that learners perceive pronunciation to be an important component of language learning, but still, they do not know what to do to improve their pronunciation skills (Darcy, 2018). For example, they may be focusing on improving accuracy in the production of L2 vowels and consonants while neglecting a focus on improved timing and speech rhythm, or they may be unaware of the importance of linking phenomena in connected speech. For many teachers, pronunciation represents a teaching challenge too. They do not know what aspects of pronunciation they should be teaching, what the main pronunciation targets should be, or when and how they should be teaching them (Darcy et al., 2012). In addition, teachers may lack knowledge of phonetics and the sound systems of the learners' first (or native) languages (L1s) and consequently they may be unable to establish pronunciation teaching priorities according to different proficiency

levels or the relative functional load of the pronunciation targets that matter the most for learners to achieve high levels of intelligibility in the L2. Even when they have such knowledge, teachers find it methodologically challenging to effectively integrate pronunciation instruction into communicative language teaching (CLT) (Darcy & Rocca, 2022; Mora-Plaza, 2023), to motivate students to work on pronunciation, and to implement methods and techniques to help them with pronunciation (Huensch, 2019).

Several features of current foreign language classroom instruction make it particularly challenging for teachers to implement efficient pronunciation instruction methods and for researchers to identify individual and context-related predictors of pronunciation learning success in this context. Two such features are the quantity and quality of INPUT learners typically receive in classroom instruction and the scarce opportunities typical meaning-focused activities provide for developing learners' AWARENESS of their L2 pronunciation learning needs:

Input. One important limitation for L2 pronunciation development in instructed SLA is the usual poor input conditions under which acquisition develops (Muñoz, 2014). Learners do not only have little opportunity for meaningful communicative interaction in the L2, the oral interactions that take place with peers and teachers are often L1-accented. Such input conditions severely limit exposure to the range of contrasting L2 sounds and features that would be necessary to guarantee the development of a fully functional L2 phonology. One undesirable consequence of such poor input conditions is that learners' vocabulary typically grows with an under-developed phonology and before phonetic contrasts have been acquired. Although L2 learners' larger vocabulary size is generally associated with more accurate pronunciation and a more precise encoding of difficult L2 phonetic contrasts (Daidone & Darcy, 2021; Llompart, 2021), vocabulary growth itself might hinder phonological development under learning conditions of extensive exposure to L1-accented input and little exposure to authentic L2 input, as imprecise ('fuzzy') phonolexical representations might get more strongly entrenched in the L2 mental lexicon through L2 vocabulary use (Darcy et al., 2025).

Awareness of pronunciation. An inherent feature of CLT is that speaking tasks are naturally meaning-oriented and, logically, pronunciation-unfocused. Consequently, the teaching focus tends to be on vocabulary and grammar because these linguistic domains, but not pronunciation, are more essential in achieving task goals (Mora-Plaza, 2023), especially at lower proficiency levels. As a consequence, learners have few opportunities to develop awareness of their pronunciation learning needs and are unlikely to identify their own pronunciation difficulties and develop strategies to overcome them.

Although much L2 learning takes place in foreign language learning environments and mainly in the classroom, current well-established models of L2 speech learning such as PAM-L2 (Best & Tyler, 2007) and SLM-r (Flege & Bohn, 2021) aim to account for L2 phonological acquisition in naturalistic language learning environments where the L2 is predominantly used and have not had much to say about how phonological development occurs in instructed SLA. By L2 phonological acquisition, these models essentially refer to the process whereby learners develop perceptual phonological categories for L2 sounds that are distinct from those of their native language and acquire the ability to make use of L2-specific features in the distinction of phonologically contrastive L2 sounds (e.g. /æ/-/ʌ/ as in *cap-cup*). Whether the predictions of these models holds under poor input conditions, such as foreign language classroom instruction, still remains an empirical question. For example, whereas in naturalistic environments, the quality and quantity of input received, the age at which immersion begins, length of residence, and the amount of L1 and L2 use have been shown to account for a large proportion of the variance in L2 speech learning, in much classroom instruction poor input conditions make phonological acquisition extremely challenging, with only very modest gains typically being reported. Consequently, the factors that explain phonological acquisition in immersion settings are unlikely to fully explain phonological acquisition in instructed SLA. This situation may change with current extensive access to online sources of spoken L2 input via media or immersive experiences provided by teachers. Researchers have generally assumed that for the models' predictions to

hold in this classroom contexts, instruction should be enriched in terms of input exposure, increased awareness of cross-language phonetic differences, and support for perceptual learning at early stages (Piske, 2007; Tyler, 2019).

Beyond the limitations of the learning conditions in instructed SLA outlined above, other factors make pronunciation learning inherently challenging. One such factor is the nature of our perceptual systems and how they cope with the acquisition of a new phonology. L2 sounds are initially perceived in terms of our L1 phonetic categories because our perceptual system is attuned to the L1 and will filter out acoustic information that is irrelevant for efficient L1 phonological processing (i.e. L1-based processing). However, such acoustic information may be critical for efficient L2 phonological processing. For example, Spanish learners of English will not attend to vowel duration differences English speakers use as a cue to voicing in word-final position (*plays* /pleɪz/ realized as [pleɪz] vs *place* /pleɪs/ realized as [pleɪs]) as there are no voiced obstruent consonants in word-final position in Spanish. Consequently, learners may not be able to perceive L2 sounds and L2-specific speech processes as faithfully as they would in their L1 (a phenomenon known as ‘perceptual foreign accent’), and they may fail to focus on the relevant L2-specific phonetic cues. Similarly, the English vowels in *cap* /kæp/ and *cup* /kʌp/ may both be perceptually assimilated to the only Spanish low central vowel category /a/, which may lead to the development of imprecise phonolexical representations. Thus, both /kæp/ and /kʌp/ might have the phonological form /kʌp/ in the learners’ mental lexicon, which may lead to both *cap* and *cup* being mispronounced undistinguishably as /kʌp/.

The consequences of the pronunciation learning challenges outlined above may have far-reaching consequences in terms of overall language development. For example, students who speak the L2 with poor intelligibility may feel that their pronunciation prevents them from communicating effectively and fluently in the L2, which may demotivate them and prevent them from seeking opportunities for interacting with other speakers or may avoid getting actively engaged in classroom speaking tasks. Rather than opportunities for language learning, speaking tasks may then be perceived as sources of language speaking anxiety interfering with pronunciation learning (Baran-Łucarz & Lee, 2021). Still, learners’ pronunciation learning goals may vary a lot depending on their overall language learning objectives. Whereas many language learners aim at acquiring a near native-like pronunciation of the L2 because their aim is to communicate with native speakers of the L2 (e.g. Italian, Russian, Japanese), the primary aim of many learners of English is to become fluent English users for intercultural communication. For them, English has the status of a ‘lingua franca’ they use for general communication purposes with other non-native speakers of English (Jenkins, 2006). For such learners, the presence of an L1-accent in their English is not a reason for concern as long as it is not seriously detrimental to intelligibility. For them, the ability to perceptually accommodate to a variety of non-native accents of English for fluent communication is more essential than speaking the L2 in a native-like manner. However, there may be a minority of learners acquiring English with the primary aim of communicating with native English speakers, or with the aim of becoming English language teachers, who may prefer to acquire a pronunciation of English that closely resembles the standard pronunciation of native speakers of some widely taught variety of English (e.g. General American English or Standard Southern British English) representing a large percentage of learners’ media consumption in the L2.

Aligned with many learners’ goals in pronunciation learning, the field of L2 pronunciation pedagogy has experienced an important paradigm shift from an emphasis on nativelikeness to an emphasis on speech intelligibility and comprehensibility as a more useful and realistic goal of pronunciation instruction (Levis, 2020). However, pedagogical practices that integrate a focus on speech comprehensibility development into CLT are scarce and empirically under-researched. Although the intelligibility principle for L2 pronunciation advocates for abandoning the nativeness principle (where L2 learners’ goal was to acquire native-like pronunciation) to embrace intelligible pronunciation (and comprehensibility) as the goal of L2 pronunciation teaching and learning, in general, ideologies of nativeness continue to be pervasive (Jeong & Lindemann, 2025). Learners vary in how

willing they may be to assume an L2 speaker identity that can speak the L2 without an L1 accent (Szyszka, 2022), which may affect the pronunciation learning strategies and goals they set for themselves (Sardegna, 2022). Similarly, although non-native speaking pronunciation teachers may idealize nativeness while relying on their knowledge of L2 phonetics and phonology to obtain legitimacy as pronunciation instructors (Gordon & Barrantes-Elizondo, 2024), the goals of pronunciation instruction and the focus on specific pronunciation targets may largely depend on teacher cognitions and identity related to pronunciation instruction (Kochem, 2022). Students' and teachers' cognitions and ideologies about L2 pronunciation teaching and learning are beyond the scope of the current paper, but they should be considered in relation to the effectiveness of the pronunciation instruction methods and training techniques described below. While some approaches to pronunciation instruction may be apt to develop L2 pronunciation from both the perspectives of intelligibility (actual understanding of speech) and nativelikeness (sounding like a native speaker), some pronunciation training techniques may be more effective than others at improving overall speech comprehensibility (ease of understanding).

Helping learners overcome the challenges associated with pronunciation learning in CLT through classroom instruction would need a thorough understanding of the pronunciation learning difficulties of specific L1-L2 combinations and learner groups, identifying the most relevant (i.e. high functional load) pronunciation targets, and implementing approaches to L2 pronunciation instruction and training methods and techniques whose effectiveness have been proven by empirical research. Given that learners' individual differences (e.g. experiential, emotional, aptitude-related) are likely to interact with one another and with teaching methods and learning contexts in complex ways to modulate individual pronunciation development (Mora, 2022), they also need to be considered when researching the effectiveness of pronunciation instruction methods and pronunciation teaching, training, and assessment.

2. Integrating pronunciation instruction into CLT: Empirical findings on pronunciation instruction methods

Several meta-analytical studies have shown that pronunciation instruction is effective at improving various segmental and suprasegmental features of L2 pronunciation (e.g. Lee et al., 2015; Thomson & Derwing, 2015), even when such techniques and methods are not integrated into CLT. Pronunciation instruction, including explicit pronunciation instruction (Gordon et al., 2013), appears to be especially effective when form-focused, when learners' attention is explicitly directed to L2-specific features, focused practice and corrective feedback is provided, the focus is on phonetic features that have an impact on intelligibility, and assessment methods elicit controlled speech that is measured acoustically (Saito & Plonsky, 2019). In addition, pronunciation research in instructed SLA is extremely varied with respect to intervention outcomes, which are largely dependent on the type (explicit, form-focused, communicative) and length of instruction, the pronunciation targets the instruction focuses on (specific segmental or suprasegmental features vs global pronunciation proficiency dimensions like comprehensibility and accentedness), and the task types used to practice pronunciation (Crowther & Loewen, 2025). The complexity of the overall emerging picture calls for further research on pronunciation instruction methods. However, research investigating the effectiveness of integrating a focus on phonetic form (FoPF) into CLT is currently scarce and findings are far from conclusive.

There are currently two main approaches to pronunciation instruction that aim at integrating pronunciation instruction into communicative language teaching. One approach advocates for combining explicit instruction with communicative tasks (Darcy et al., 2021; Gordon, 2021; Gordon & Darcy, 2022), as the addition of communicative activities focusing on the target phonetic features previously taught has been shown to provide a learning advantage. For example,

Darcy & Rocca (2022) compared two types of pronunciation instruction, explicit only versus explicit + communicative. Pronunciation instruction, which focused on suprasegmentals, was integrated into a seven-week oral communication course (110 instruction hours). A control group that did not integrate a specific focus on pronunciation provided baseline data. By comparing scores before and after the intervention, the researchers tested gains in comprehensibility and perceived errors (stress and vowel reduction) on speech samples obtained through a read aloud task, a monologic spontaneous speaking task, and an interactive group discussion task (information gap task). They found improvement in comprehensibility when phonetic instruction was integrated in both the explicit only and the explicit + communicative groups, but not in the pronunciation-unfocused group. However, the group with a communicative component did better in the group discussion task than the explicit group, whereas the latter obtained higher gains in the read aloud task and both groups improved only modestly in the monologic speaking task. Much more classroom research of this kind is needed to gain a better understanding of how and to what extent integrating a focus on pronunciation into CLT is possible and effective.

Another approach is task-based pronunciation teaching (TBPT). This approach applies the principles of task-based language teaching (TBLT) to L2 pronunciation teaching and learning (Mora & Levkina, 2017; Mora & Mora-Plaza, 2023; Mora-Plaza, 2023; Solon et al., 2017; Xu et al., 2024). The central idea is to promote attention to linguistic form during communicative task performance by manipulating task design features such as task complexity. This approach is grounded in TBLT research showing that increasing the cognitive complexity of a task enhances attention to linguistic form, thus promoting oral production that is lexically and grammatically more complex (and often more accurate), enhancing linguistic development. Research on TBPT has investigated whether this holds true for pronunciation. Some studies have found an increase in perceptual or productive segmental accuracy resulting from increased cognitive complexity when pronunciation practice is provided through meaning-oriented pronunciation-focused tasks, especially when the target phonetic feature is made essential for task completion. For example, Mora-Plaza (2023) implemented a seven-week pedagogic intervention with 92 L1-Spanish EFL school learners (aged 16) around the topic of a trip consisting of 20 pronunciation-focused dyadic problem-solving communicative tasks (e.g. deciding on destinations and visits, choosing which objects to take on the trip) based on minimal-pair words targeting challenging phonological contrasts (/i:/-/ɪ/, /æ/-/ʌ/). Learners, who were assigned to either a simple or a complex task condition (or a control group), engaged in pre-tasks that familiarized them with the target word forms and in post-tasks that aimed to consolidate the target vowel contrasts through self-assessment before and after the communicative tasks, respectively. The key strategy in the task design of the study is that target forms are practiced during meaningful communication between peers, because this is what is meant to drive acquisition. She found both experimental learner groups to improve significantly in the perception and production of the /i:/-/ɪ/ and /æ/-/ʌ/ contrasts, but those in the complex task condition obtained higher retention rates, suggesting that, in these kinds of pronunciation-focused tasks, increased task complexity may generate more robust learning gains in pronunciation. However, in pronunciation-unfocused tasks increased task complexity can be detrimental to speaking fluency and pronunciation accuracy in terms of accentedness, rhythm, and intonation (Crowther et al., 2018). This raises the question of whether learners can actually attend to pronunciation in meaning-oriented pronunciation-unfocused tasks and, especially in cognitively complex tasks, whether enhanced attention to form generated by increased complexity would be detrimental to pronunciation by favouring increased attention to grammar and lexis. Such a question has important implications for pedagogic approaches that aim to integrate pronunciation instruction into task-based CLT, especially in terms of communicative task design and implementation.

Empirical research has yet to determine the extent to which there may exist attentional trade-offs between pronunciation and lexico-grammar during meaning-based activities with detrimental

effects for pronunciation, or the extent to which such trade-offs can be manipulated through task design to enhance attention to phonetic form. Mora-Plaza et al. (2024) is a first attempt at investigating this question. This study manipulated task complexity and its effects on specific features of L2 English pronunciation in the speech of Spanish EFL learners (VOT, vowel production accuracy, and speech rhythm) in a problem-solving pronunciation-unfocused monologic speaking task (an adaptation of the dinner table task, available from <http://sla-speech-tools.com>, Mora Plaza et al., 2022). Pronunciation accuracy was gauged through acoustic measurements of laryngeal timing (voice onset time, VOT), vowel contrastiveness and nativelikeness (Mahalanobis distances), and native speakers' ratings of comprehensibility and accentedness. Results revealed detrimental effects of increased task complexity on the production of oral stops and speech accentedness; however, no consistent task complexity effects were found on vowel accuracy. In a follow-up study using the same speech materials but focusing on L2 speech rhythm (Fullana et al., 2025), the oral production data were analysed using well-established rhythm metrics (%V, VarcoV, nPVI-V, and VarcoC), novel distance measures (Euclidean and Mahalanobis distance scores between pairs of rhythm metrics), and native speaker ratings of comprehensibility and accentedness. The results showed that differential task complexity effects on L2 speech rhythm were dependent on the specific rhythm metric or distance measure used, but an overall consistent clear picture did not emerge that could confirm potential detrimental effects of task complexity on L2 speech rhythm. Task complexity was only slightly detrimental to comprehensibility and accentedness.

These findings suggest that increasing task complexity in pronunciation-unfocused tasks may be detrimental for L2 pronunciation learning, such that if pronunciation-unfocused, simple (but not complex) tasks may be more likely to enhance some focus on phonetic form during communicative task performance. In pronunciation-focused communicative tasks, on the other hand, a TBPT task design that makes pronunciation targets essential for task completion may be effective in promoting pronunciation learning in the classroom. Pronunciation pedagogy would benefit substantially from empirical research investigating the effectiveness of integrated L2 pronunciation instruction approaches, currently very scarce. In addition, further research is needed in various areas of L2 pronunciation teaching and training methods, such as the extent to which high variability phonetic training (HVPT) may be effective at developing pronunciation globally, or the extent to which specific pronunciation training methods (e.g. accent imitation, shadowing) can be useful in supporting pronunciation development by enhancing learners' awareness of their pronunciation limitations and learning needs. Finally, without helping learners establish precise phonological representations for new words and helping them update the representations of already established lexical forms, L2 pronunciation training is unlikely to lead to substantial benefits in L2 speech comprehensibility and accentedness (Darcy et al., 2025). Consequently, it becomes crucial to identify pronunciation methods and techniques that hold a potential for accomplishing such goals, as well as to empirically test whether certain task-design features (for example, the use of form-meaning mapping trials when training difficult sound contrasts through HVPT) can benefit the establishment and updating of phonolexical representations more than others. Therefore, a potentially fruitful research avenue in the area of L2 pronunciation teaching and learning might be to compare and combine specific and global methods of L2 pronunciation training with a focus on the development of pronunciation at the lexical level and to examine the role of aptitude-related individual differences, such as domain-general auditory processing and cognitive factors.

3. An outline of pronunciation training techniques: Characteristics, limitations, and empirical findings

There is little empirical evidence of the positive impact of phonetic training methods on the development of comprehensible speech. Pronunciation instruction approaches like TBPT and explicit

instruction (+ communicative tasks) outlined above have produced promising results, but much more research is needed to determine their efficacy in learners of varying proficiency levels, in different learning contexts, and on a wider range of segmental and suprasegmental phonetic targets. Similarly, the efficacy of pronunciation training techniques available to teachers for in and out of class use is largely under-researched. Such techniques include, but are not limited to, the following:

High variability phonetic training (HVPT): In HVPT learners are typically perceptually trained on difficult L2 vocalic or consonantal contrasts (e.g. /r/-/l/ for Japanese learners of English or /æ/-/ʌ/ for Spanish learners of English) through exposure to highly variable stimuli (different speakers and phonetic environments) via identification or discrimination tasks with feedback (Thomson, 2018). High variability phonetic training has been found to produce robust gains in L2 speech perception and production generalizable to new speakers and contexts (Sakai & Moorman, 2018; Uchiyama et al., 2024, 2025) and has been successfully adapted to implicit gaming paradigms (Saito et al., 2022). A potential pedagogical limitation of HVPT, however, is the fact that it is more suitable for individual pronunciation training than integrated classroom practice.

Shadowing: Shadowing involves repeating what one hears as simultaneously and accurately as possible. When shadowing speech presented auditorily, learners are forced to process L2 sound units, sequences, and phonological features while avoiding a focus on meaning during L2 speech input processing. Shadowing has been shown to improve speech perception and production and listening skills (Foote & McDonough, 2017; Hamada, 2019). This technique is highly adaptable and has been used in various forms to train phonological and bottom-up processing skills (Hamada & Suzuki, 2024).

Embodied pronunciation training (EPT): EPT involves the use of multimodal exposure through gestures to visually support pronunciation training by linking auditory perception with visual actions to enhance the acquisition of segmental and suprasegmental features of speech (Chan, 2018). This technique, which may involve hand clapping, hand movements representing tones or intonational contours, or a fist-to-open-hand gesture representing the release burst of aspirated stops, among other gestures, has been shown to result in pronunciation benefits in specific (segmental and suprasegmental accuracy) and global (comprehensibility, accentedness) speech dimensions (Baills et al., 2022; Xi et al., 2024).

Multimodal pronunciation training through L2-captioned video (CaptV): Captioned video is a useful tool to train the simultaneous processing of L2 auditory input with the visual support of onscreen text and background action. Benefits of exposure to this kind of audiovisual input have been widely attested for listening comprehension and for vocabulary and grammar development (Muñoz, 2022), but also for speech segmentation (Charles & Trenkic, 2015) and speech perception (Mitterer & McQueen, 2009). Although still an under-researched method for pronunciation learning, several studies suggest it has the potential to generate benefits at the segmental processing level and in speech perception (Galimberti et al., 2023; Scheffler & Baranowska, 2023; Wisniewska & Mora, 2020) and production (Hutchinson & Dmitrieva, 2022). Given the difficulty in directing learners' attention to phonetic form under such attention-demanding rich input conditions, studies have used question prompts (Wisniewska & Mora, 2020), input-based activities (Galimberti et al., 2025), or input enhancement (Azpilicueta-Martínez & Ocariz-Tejada, 2025; Fouz-González & Mora, 2025) to bring target pronunciation features to the attentional foreground during viewing.

L2 accent imitation in L1 (L2AIL1): L2AIL1 involves mimicking the accent of a speaker of the L2 one is learning while speaking one's L1, usually after having been exposed to such an accent through listening materials. For example, a Spanish learner of English would imitate an English speaker speaking Spanish with an English accent. This technique has been found useful in developing awareness of cross-language phonetic differences resulting in pronunciation improvement at the segmental level (Henderson & Rojczyk, 2023; Mora et al., 2014; Rojczyk, 2015).

Pronunciation self- and peer-assessment (SelfAss): Self- and peer-assessment of pronunciation are useful tools to raise learners' awareness of their own pronunciation difficulties and help them identify

the features they need to improve on (Isbell & Lee, 2022; Saito et al., 2020). Pronunciation self-assessments are typically compared to native listeners' assessments to obtain measures of under- and over-estimation of learners' perceived pronunciation ability (Strachan et al., 2019; Trofimovich et al., 2016). Training self-assessment may therefore be a way to achieve a better alignment between assessment by learners and native listeners, which may have positive consequences for L2 pronunciation development (Tsunemoto et al., 2022).

One way to orient research on the effectiveness of the pronunciation training techniques briefly outlined above is to determine their appropriateness in terms of a set of critical features (see Table 1 below), such as learners' proficiency level (low, intermediate, advanced), learning context (classroom vs self-study), the processing level the training involves (pre-lexical or lexical), the pronunciation targets they focus on (vowels, consonants, suprasegmentals), the training modality (primarily perception or production), the training focus (L2 features or contrasting L2-L1 features), or the specific (segmental contrasts, prosodic features) or global dimensions (intelligibility, comprehensibility, accentedness, fluency, listening) they are most likely to have an effect on (see Saito & Plonsky, 2019). Empirical studies are needed for every intersection between each one of the training techniques and the different task features in Table 1. For example, HVPT has been widely used to perceptually train L2 segmental contrasts (using lexical and non-lexical materials) in intermediate and advanced adult learners, but very little research has investigated its effectiveness in training the perception of cross-language phonetic differences (Cebrian et al., 2025) or in training young children (Brekelmans et al., 2024).

The pronunciation training techniques in Table 1, and many others, are likely to develop dramatically in the near future through technological advances (Fouz-González, *in press*; O'Brien et al., 2018), especially through the implementation of computer- and mobile-assisted pronunciation training (CAPT/MAPT), which has been found to be effective in promoting pronunciation development (Stoughton & Kang, 2024), but also through the use of artificial intelligence (AI) and automated speech recognition (ASR) systems (Dai & Wu, 2023). Undoubtedly, both the design and implementation of the pronunciation training techniques described above will benefit greatly from technological advances, including the use of ASR to train L2 pronunciation and AI-generated speaker voices to generate speech stimuli for training and testing.

Another important, still largely under-researched aspect of L2 pronunciation training interventions (whether focusing on specific or global dimensions of pronunciation proficiency) is the extent to which they can have an impact on the processes that modify learners' developing phonolexical representations and whether such an impact can have positive consequences for L2 speech production (John & Frasnelli, 2022) or they can contribute to enhance the spreading of perception-based phonolexical updates to the entire L2 lexicon and to spontaneous L2 speech production (Rocca et al., 2025). At present, the mechanisms that underlie the updating of phonolexical representations in the mental lexicon are not fully understood and therefore it is currently not possible to implement specific training methods that would be maximally efficient at changing the imprecise pronunciation of words learners have established in their L2 lexicons.

In addition, there is large variability seen in L2 learners' phonological acquisition, in general and in instructed SLA in particular, in terms of pronunciation instruction and training gains and overall attainment levels. Understanding the multi-componential nature of the sources that underlie inter-learner variability in phonological acquisition is fundamental in adapting methods to learners' pronunciation learning needs and their individual cognitive and emotional profiles (Mora, 2022). Research into L2 pronunciation has only recently begun to investigate the complex role of individual differences in L2 pronunciation learning in instructed SLA. The preliminary picture emerging suggests that individual factors interact with one another in complex ways to enhance or constrain pronunciation development and points to the need of further research in this area, as outlined in the following section.

Table 1. Task features to consider when assessing effectiveness of pronunciation training techniques

Training Techniques	Task Features						
	Proficiency	Processing	Target	Modality	Focus	Context	Dimension
HVPT	Intermediate Advanced	Pre-lexical Lexical	Segments	Perception (Production)	(L2-L1) L2	(Classroom) Self-study	Acc
Shadowing	Intermediate Advanced	Pre-lexical Lexical	Segments Prosody	(Perception) Production	L2	(Classroom) Self-study	Acc Int Comp Flu List
EPT	Low	Pre-lexical Lexical	Segments Prosody	Perception mProduc- tion	L2	Classroom (Self-study)	Acc Comp
CaptV	Intermediate	Lexical	Segments	Perception (Production)	L2	(Classroom) Self-study	Acc Int Comp
L2AIL1	Low Intermediate	Lexical	Segments	Production	L2-L1	Classroom Self-study	Acc Comp
SelfAss	Intermediate Advanced	Lexical	Segments Prosody	Perception Production	L2	Classroom Self-study	Acc Comp

Note: Acc = accentedness, Comp = comprehensibility, Flu = fluency, Int = intelligibility, List = listening.

4. Domain-general auditory processing skills and cognitive and emotional individual differences in L2 pronunciation learning

Sources of individual differences in L2 pronunciation learning, which include experience-related factors linked to overall L2 proficiency (e.g. vocabulary size, amount of L2 use), socio-psychological emotional variables linked to pronunciation learning (e.g. motivation, anxiety, enjoyment), domain-general auditory processing skills (e.g. frequency discrimination), and cognitive and aptitude-related factors (e.g. memory, inhibition, and attention), interact with one another to contribute differentially to L2 pronunciation learning as a function of individual learner profiles, learning contexts, and pronunciation teaching approaches. One way to think about individual differences factors is to conceptualize them as cue-enhancement devices. For example, good attention control is likely to help learners pay attention to relevant L2-specific phonetic cues during L2 speech processing. Another way to conceptualize them is in terms of a bottleneck that determines how much learners can benefit from auditory input. Several factors might play this role in L2 speech learning, but they would also vary as to their relative weight at an individual level. For example, for someone receiving a particular type of input (e.g. three weekly hours of classroom instruction + regular weekly exposure to L2 captioned video) through a given type of instruction (form-focused grammar-centred CLT), auditory processing skills may function as a bottleneck and then other factors (e.g. L2 proficiency, anxiety, attention control, or motivation) may further constrain input processing to varying degrees. Saito's (2023) auditory precision hypothesis-L2 advocates for the fundamental role of domain-general auditory processing in L2 speech acquisition in allowing learners to maximally benefit from input exposure and pronunciation training, thus leading to more advanced L2 pronunciation proficiency. Within this view, auditory processing skills are supported by cognitive attentional mechanisms that allow learners to efficiently direct their attention towards relevant L2-specific phonetic features and by integration mechanisms that allow them to transform perceptual representations into motor actions generating L2 speech output (Saito et al., 2024). The interaction between auditory processing skills and attentional and integration skills are partly determined by how learners differ from one another in such skills. In addition, the interaction between instructional conditions, input exposure, and cognitive and auditory processing skills are likely to shape L2 learners' phonolexical representations, and eventually pronunciation accuracy during L2 use. That is, individual differences will affect individual learners differently in shaping their representations and ultimately their pronunciation development. For the learner model represented in Fig. 1 below, domain-general auditory processing skills appear to be the primary source of individual differences in L2 speech learning (1), constraining how much the learner can benefit from the L2 speech input and exposure received through pronunciation instruction and learning, followed by attention control (2), motivation (3), emotions such as speaking anxiety (4), and proficiency factors like vocabulary size (5), in decreasing order of importance. The specific combination of individual differences factors and their relative weight will partly determine the nature and shape the learners' phonetic, phonological, and phonolexical representations, which will eventually be reflected in the learners' speech output during L2 use.

Whether certain factors may stand out as more fundamental (for most learners in general) in constraining how much learners can benefit from L2 exposure and pronunciation training and instruction, or the extent to which different factors play a stronger role for different learners or learning environments, still remains an empirical question in need of further research. For example, some researchers have hypothesized that it is domain-general auditory processing skills and cognitive factors that are most fundamental (e.g. Saito, 2023). Investigating the joint contribution of auditory processing and cognitive skills in L2 pronunciation learning may help clarify some of the mixed findings on the role of cognitive individual differences in L2 speech learning (e.g. Saito et al., 2024). From a research perspective, the key difficulty in determining the role of individual factors in pronunciation development in instructed SLA lies in the possibility of any given factor playing a role for specific learner groups and learning contexts, but not for others. Critical methodological features of research

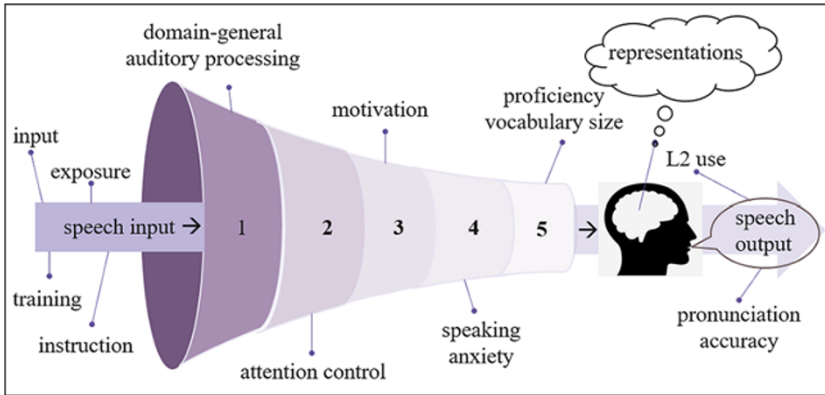


Figure 1. Model of a learner with individual differences factors in order of importance (1–5).

design such as sample size, whether the data is longitudinal or cross-sectional, or whether several factors are tested on the same participants within the same study, will also affect the extent to which the contribution of individual differences factors can reliably be measured. For example, in a recent study, Saito et al. (2025) investigated the relationship between individual differences in motivation (ideal self and ought-to self), emotions (enjoyment and anxiety), and L2 speech learning longitudinally (three testing times in 1.5 years, a rarity in individual differences research) among 121 Japanese EFL high school students. Unlike findings in previous cross-sectional research pointing to its detrimental effect on language learning, anxiety did not surface as a significant predictor of L2 speech learning, whereas both motivation and enjoyment predicted increased classroom practice and led to significant speech learning. The authors point out that anxiety may play less of a role longitudinally than it does when its effects are measured cross-sectionally, which highlights the importance of conducting more longitudinal research on individual differences (Nagle, 2023).

Research on the role of attention control in L2 speech learning offers another example of mixed findings related to learner populations and pronunciation targets (perception and production of vowels and consonants). In two related studies on the role of inhibitory control and attention control with instructed L2 learners (see Table 2), we obtained inconsistent results related to learner populations and testing targets despite using the same testing instruments for cognitive skills (a retrieval-induced forgetting task for inhibition and an alternating runs attention-switching task for attention) and pronunciation proficiency (a categorical ABX discrimination task in perception and a delayed sentence repetition task in production).

In Study 1 (Darcy et al., 2016; Mora & Darcy, 2023) inhibition and attention were tested bidirectionally on comparable populations of English learners of Spanish and Spanish learners of English. Inhibitory control explained a significant 17.6% of the variance in vowel perception, but not on consonant perception, whereas in production it explained a non-significant 11.6% in consonant but not vowel production. On the other hand, attention control was found to explain some (non-significant) variance only in vowel production (11.2 %). In Study 2 (Darcy & Mora, 2016; Mora & Darcy, 2016), two groups of learners, monolingual Spanish speakers and bilingual Spanish-Catalan speakers, were tested on their inhibitory and attention control skills and their L2 English perception and production skills as in Study 1. However, with this population, inhibition was found to explain a significant 46.2% of the variance in vowel and consonant perception (not in production), but only in the monolingual group (not in the comparable bilingual group). On the other hand, attention control explained a significant amount of variance in vowel and consonant production (but not perception) and only for the monolingual group. These results suggest that inhibition and attention contribute differently to vowels and consonants, to perception and production, and to monolingual and bilingual populations.

Table 2. Summary of outcomes of individual differences studies on inhibition and attention

Predictors	Study 1		Study 2	
	English learners of Spanish Spanish learners of English		Monolingual EFL learners (Mon) Bilingual EFL learners (Bil)	
	Perception	Production	Perception	Production
	ABX (Vs & Cs)	DSR (Vs & Cs)	ABX (Vs & Cs)	DSR (Vs & Cs)
Inhibition (retrieval-induced forgetting)	17.6%* ✓ (Vs not Cs)	11.6% [†] ✗ (Cs not Vs)	46.2%* ✓ (Mon not Bil)	< 1% ✗ (Mon & Bil)
Attention (attention switching)	< 1% ✗ (Vs)	< 11.2% [†] ✗ (Vs)	< 1% ✗ (Mon & Bil)	44.1%* vs 15.1% ✓ (Mon not Bil)

*indicates significance, [†]indicates marginal significance.

Interestingly, Huensch (2024) replicated the first of these studies on inhibitory control and added two more inhibition tasks (Simon, Stroop) with L1 English learners of Spanish and found that cognitive measures only explained a small non-significant amount of variance (1%–6%) in the perception and production of vowels and consonants.

To summarize, previous and current research on individual differences in L2 speech learning, especially in instructed SLA, suggests that individual variation is pervasive throughout and remains a very challenging research endeavour. Experiential, affective, cognitive and auditory processing factors are likely to interact with one another differently in different learners according to individual learning styles, learning contexts, instructional methods, and training conditions and tasks, making the explanation of the sources of inter-learner variation in L2 pronunciation proficiency extremely challenging.

5. Conclusions

This article has presented an overview of challenging research topics in L2 pronunciation learning and teaching in instructed SLA with the aim of highlighting fruitful avenues for future research. One such challenge is finding out whether the principles and predictions of current models of L2 speech learning developed to account for phonological development in naturalistic environments, such as PAM-L2 and SLM-r, can be applied to and hold for instructed SLA. To this end, research should investigate the long-term effects of enriching input conditions and training cross-language phonetic differences and L2 segmental contrasts at an early stage. Still, a focus on pronunciation is not easy to integrate into communicative approaches to language teaching that prioritize meaning-focused activities. Attempts at integrating a focus on phonetic form into CLT (e.g. TBPT) are currently scarce and under-researched but constitute a fruitful approach to L2 pronunciation teaching and learning that should be explored and developed further, both from pedagogical and research perspectives.

A second important challenge involves adjusting instructional methods to reflect the recent paradigm shift from an emphasis on accent reduction and nativelikeness to an emphasis on the development of comprehensible speech and intelligible pronunciation. Although traditional explicit form-focused pronunciation instruction as well as approaches that combine pronunciation instruction with communicative practice lead to benefits in L2 pronunciation, there is still little empirical evidence of their effectiveness to develop global dimensions of L2 pronunciation proficiency (e.g. comprehensibility). Recent technological advances (AI, ASR) in combination with rich input-based pronunciation training techniques (shadowing, textually-enhanced captioned video, accent imitation) may provide new ways of training pronunciation globally to effect larger impacts on the development of overall speech intelligibility.

Finally, a third challenge with important pedagogical implications involves gaining a better understanding of the sources of inter-learner variation in L2 pronunciation performance, development, and attainment. Research-informed knowledge of the relative contribution of different factors (auditory processing skills, cognitive individual differences in attention control, motivation and emotional variables) to L2 speech learning in relation to the characteristics of specific learner populations under different instructional methods and in a variety of learning contexts would be very valuable in developing tasks and training methods tailored to match their pronunciation learning needs. I hope that the three research challenges outlined above, among others, pave the way for exciting new research avenues in L2 pronunciation teaching and learning.

References

- Azpilicueta-Martínez, R., & Ocariz-Tejada, N. (2025). The impact of audio-synchronised keyword captions on young CLIL learners' foreign language pronunciation. *System* 133, 103730. Advance online publication. <https://doi.org/10.1016/j.system.2025.103730>
- Baills, F., Alazard, C., & Prieto, P. (2022). Embodied prosodic training helps improve accentedness and suprasegmental accuracy. *Applied Linguistics*, 43(4), 776–804. <https://doi.org/10.1093/applin/amac010>
- Baran-Lucarz, M., & Lee, J.-H. (2021). Selected determinants of pronunciation anxiety. *International Journal of English Studies*, 21(1), 93–113. <https://doi.org/10.6018/ijes.426411>
- Best, C., & Tyler, M. (2007). Non-native and second language speech perception. In O.-S. Bohn, & M. J. Munro (Eds.), *Language experience in second language speech learning* (pp. 15–34). John Benjamins.
- Brekelmans, G., Evans, B. G., & Wonnacott, E. (2024). Training child learners on nonnative vowel contrasts with phonetic training: The role of task and variability. *Language Learning*, 468–505. Advance online publication. <https://doi.org/10.1111/lang.12677>
- Cebrian, J., Cortés, S., Gavalda, N., Mora Plaza, I., Gorba, C., De Witte, W., & Carlet, A. (2025, May 9). *Challenges in HVPT experiment design revisited: Training learners to distinguish L2 and L1 categories (II)*. Paper presented at the Segundas Jornadas Metodológicas sobre Adquisición y Fonética de la L2, University of the Basque Country, Vitoria, Spain.
- Chan, M. J. (2018). Embodied pronunciation learning: Research and practice. *The CATESOL Journal*, 30(1), 47–68. <https://doi.org/10.5070/B5.35964>
- Charles, T. J., & Trenkic, D. (2015). Speech segmentation in a second language: The role of bimodal input. In Y. Gambier, A. Caimi, & C. Mariotti (Eds.), *Subtitles and language learning: Principles, strategies and practical experiences* (pp. 173–198). Peter Lang.
- Crowther, D., & Loewen, S. (2025). Instructed second language acquisition and second language pronunciation. *Language Learning*, 30–54. Advance online publication. <https://doi.org/10.1111/lang.12716>
- Crowther, D., Trofimovich, P., Saito, K., & Isaacs, T. (2018). Linguistic dimensions of L2 accentedness and comprehensibility vary across speaking tasks. *Studies in Second Language Acquisition*, 40(2), 443–457. <https://doi.org/10.1017/S027226311700016X>
- Dai, Y., & Wu, Z. (2023). Mobile-assisted pronunciation learning with feedback from peers and/or automatic speech recognition: A mixed-methods study. *Computer Assisted Language Learning*, 36(5–6), 861–884. <https://doi.org/10.1080/09588221.2021.1952272>
- Daidone, D., & Darcy, I. (2021). Vocabulary size is a key factor in predicting second language lexical encoding accuracy. *Frontiers in Psychology*, 12, 688356. <https://doi.org/10.3389/fpsyg.2021.688356>
- Darcy, I. (2018). Powerful and effective pronunciation instruction: How can we achieve it? *The CATESOL Journal*, 30(1), 13–45. <https://doi.org/10.1075/bpa.3.12dar>
- Darcy, I., Ewert, D., & Lidster, R. (2012). Bringing pronunciation instruction back into the classroom: An ESL teachers' pronunciation toolbox. In J. Levis, & K. LeVelle (Eds.), *Proceedings of the 3rd pronunciation in second language learning and teaching conference* (pp. 93–108). Iowa State University. <https://hdl.handle.net/2022/22866>
- Darcy, I., Llompарт, M., Hayes-Harb, R., Mora, J. C., Adrian, M., Cook, S., & Ernestus, M. (2025). Phonological processing and the L2 mental lexicon: Looking back and moving forward. *Studies in Second Language Acquisition*, 47(1), 361–387. <https://doi.org/10.1017/S0272263124000482>
- Darcy, I., & Mora, J. C. (2016). Executive control and phonological processing in language acquisition: The role of early bilingual experience in learning an additional language. In G. Granena, D. O. Jackson, & Y. Yilmaz (Eds.), *Cognitive individual differences in L2 processing and acquisition* (pp. 249–277). John Benjamins. <https://doi.org/10.1075/bpa.3.12dar>
- Darcy, I., Mora, J. C., & Daidone, D. (2016). The role of inhibitory control in second language phonological processing. *Language Learning*, 66(4), 741–773. <https://doi.org/10.1111/lang.12161>
- Darcy, I., & Rocca, B. (2022). Comprehensibility improvements in integrated pronunciation instruction: A comparison of instructional methods and task effects. *Journal of Second Language Pronunciation*, 8(3), 328–362. <https://doi.org/10.1075/jslp.21035.dar>

- Darcy, I., Rocca, B., & Hancock, Z. (2021). A window into the classroom: How teachers integrate pronunciation instruction. *RELC Journal*, 52(1), 110–127. <https://doi.org/10.1177/0033688220964269>
- Flège, J. E., & Bohn, O.-S. (2021). The revised Speech Learning Model (SLM-r). In R. Wayland (Ed.), *Second language speech learning: Theoretical and empirical progress* (pp. 3–83). Cambridge University Press. <https://doi.org/10.1017/9781108886901.002>
- Foote, J. A., & 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>
- Fouz-González, J. (in press). *Technology-assisted pronunciation training: Bridging research and pedagogy*. University of Toronto Press.
- Fouz-González, J., & Mora, J. C. (2025). Exploring the potential of textually-enhanced captioned video to direct learners' attention to challenging sound contrasts. *Journal of Second Language Pronunciation*, 10(3), 309–338. <https://doi.org/10.1075/jslp.24043.fou>
- Fullana, N., Mora-Plaza, I., Mora, J. C., Adrian, M., & Sosa-López, G. (2025). Task complexity effects on L2 speech rhythm in spontaneous speech production. *Second Language Research*, 41(4), 621–647. <https://doi.org/10.1177/02676583241281676>
- Galimberti, V., Mora, J. C., & Gilabert, R. (2023). Audio-synchronized textual enhancement in foreign language pronunciation learning from videos. *System*, 116, 103078. <https://doi.org/10.1016/j.system.2023.103078>
- Galimberti, V., Mora, J. C., & Gilabert, R. (2025). Teaching English pronunciation with TV series: Learning outcomes from exposure to textually-enhanced video and audiovisual activities. In A. Pattemore, & F. Gesa (Eds.), *Foreign language learning from audiovisual input* (pp. 113–138). Springer. https://doi.org/10.1007/978-3-031-91001-2_6
- Gordon, J. (2021). Pronunciation and task-based instruction: Effects of a classroom intervention. *RELC Journal*, 52(1), 94–109. <https://doi.org/10.1177/0033688220986919>
- Gordon, J., & Barrantes-Elizondo, L. (2024). Idealizing nativeness vs embracing nonnativeness: A case study on L2 pronunciation teachers' identity. *Language Teaching Research. OnlineFirst*, March 3, 2024. <https://doi.org/10.1177/13621688241233840>
- Gordon, J., & Darcy, I. (2022). Teaching segmentals and suprasegmentals: Effects of explicit pronunciation instruction on comprehensibility, fluency, and accentedness. *Journal of Second Language Pronunciation*, 8(2), 168–195. <https://doi.org/10.1075/jslp.21042.gor>
- Gordon, J., Darcy, I., & Ewert, D. (2013). Pronunciation teaching and learning: Effects of explicit phonetic instruction in the L2 classroom. In J. Levis, & K. LeVelle (Eds.), *Proceedings of the 4th Pronunciation in Second Language Learning and Teaching Conference* (pp. 194–206). Iowa State University.
- Hamada, Y. (2019). Shadowing: What is it? How to use it? Where will it go? *RELC Journal*, 50(3), 386–393. <https://doi.org/10.1177/0033688218771380>
- Hamada, Y., & Suzuki, Y. (2024). Situating shadowing in the framework of deliberate practice: A guide to using 16 techniques. *RELC Journal*, 55(1), 219–227. <https://doi.org/10.1177/00336882221087508>
- Henderson, A., & Rojczyk, A. (2023). Foreign language accent imitation: Matching production with perception. In V. Sardegna, & A. Jarosz (Eds.), *English pronunciation teaching: Theory, practice and research findings* (pp. 115–133). Multilingual Matters. <https://doi.org/10.21832/9781800410503-013>
- Huensch, A. (2019). Pronunciation in foreign language classrooms: Instructors' training, classroom practices, and beliefs. *Language Teaching Research*, 23(6), 745–764. <https://doi.org/10.1177/1362168818767182>
- Huensch, A. (2024). Clarifying the role of inhibitory control in L2 phonological processing: A preregistered, close replication of Darcy et al. (2016). *Studies in Second Language Acquisition*, 46(5), 1392–1412. <https://doi.org/10.1017/S0272263124000238>
- Hutchinson, A. E., & Dmitrieva, O. (2022). Exposure to speech via foreign film and its effects on non-native vowel production and perception. *Journal of Phonetics*, 95, 101189. <https://doi.org/10.1016/j.wocn.2022.101189>
- Isbell, D. R., & Lee, J. (2022). Self-assessment of comprehensibility and accentedness in second language Korean. *Language Learning*, 72(3), 806–852. <https://doi.org/10.1111/lang.12497>
- Jenkins, J. (2006). Points of view and blind spots: ELF and SLA. *International Journal of Applied Linguistics*, 16(2), 137–162. <https://doi.org/10.1111/j.1473-4192.2006.00111.x>
- Jeong, H., & Lindemann, S. (2025). Beyond ideologies of nativeness in the intelligibility principle for L2 English pronunciation: A corpus-supported review. *System*, 103, 599. <https://doi.org/10.1016/j.system.2025.103599>
- John, P., & Frasnelli, J. (2022). On the lexical source of variable L2 phoneme production. *The Mental Lexicon*, 17(2), 239–276. <https://doi.org/10.1075/ml.22002.joh>
- Kochem, T. (2022). Exploring the connection between teacher training and teacher cognitions related to L2 pronunciation instruction. *TESOL Quarterly*, 56(4), 1136–1162. <https://doi.org/10.1002/tesq.3095>
- Lee, J., Jang, J., & Plonsky, L. (2015). The effectiveness of second language pronunciation instruction: A meta-analysis. *Applied Linguistics*, 36(3), 345–366. <https://doi.org/10.1093/applin/amu040>
- Levis, J. M. (2020). Revisiting the intelligibility and nativeness principles. *Journal of Second Language Pronunciation*, 6(3), 310–328. <https://doi.org/10.1075/jslp.20050.lev>
- Llompert, M. (2021). Phonetic categorization ability and vocabulary size contribute to the encoding of difficult second-language phonological contrasts into the lexicon. *Bilingualism: Language and Cognition*, 24(3), 481–496. <https://doi.org/10.1017/S1366728920000656>

- Mitterer, H., & McQueen, J. (2009). Foreign subtitles help but native-language subtitles harm foreign speech perception. *PLOS ONE*, 4(11), e7785. <https://doi.org/10.1371/journal.pone.0007785>
- Mora, J. C. (2022). Aptitude and individual differences. In T. M. Derwing, M. J. Munro, & R. I. Thomson (Eds.), *The Routledge handbook of second language acquisition and speaking* (pp. 68–82). Routledge. <https://doi.org/10.4324/9781003022497>
- Mora, J. C., & Darcy, I. (2016). The relationship between cognitive control and pronunciation in a second language. In T. Isaacs, & P. Trofimovich (Eds.), *Second language pronunciation assessment: Interdisciplinary perspectives* (pp. 95–120). Multilingual Matters. <https://doi.org/10.21832/ISAACS6848>
- Mora, J. C., & Darcy, I. (2023). Individual differences in attention control and the processing of phonological contrasts in a second language. *Phonetica*, 80(3–4), 153–184. <https://doi.org/10.1515/phon-2022-0020>
- Mora, J. C., & Levkina, M. (2017). Task-based pronunciation teaching and research: Key issues and future directions. *Studies in Second Language Acquisition*, 39(2), 381–399. <https://doi.org/10.1515/phon-2022-0020>
- Mora, J. C., & Mora-Plaza, I. (2023). From research in the lab to pedagogical practices in the EFL classroom: The case of task-based pronunciation teaching. *Education Sciences*, 13(10), Article 1042. <https://doi.org/10.3390/educsci13101042>
- Mora, J. C., Rochdi, Y., & Kivistö-de-souza, H. (2014). Mimicking accented speech as L2 phonological awareness. *Language Awareness*, 23(1–2), 57–75. <https://doi.org/10.1080/09658416.2013.863898>
- Mora-Plaza, I. (2023). *Task-based pronunciation teaching and learning of L2 vowels in EFL learners: Task complexity effects* (Doctoral dissertation, Universitat de Barcelona). <http://hdl.handle.net/10803/688956>
- Mora-Plaza, I., Mora, J. C., Ortega, M., & Aliaga-García, C. (2024). Is L2 pronunciation affected by increased task complexity in pronunciation-unfocused speaking tasks? *Studies in Second Language Acquisition*, 46(4), 1117–1149. <https://doi.org/10.1017/S0272263124000470>
- Mora-Plaza, I., Saito, K., Suzukida, Y., Dewaele, J.-M., & Tierney, A. (2022). Tools for second language speech research and teaching. 79(1), 1–43. <http://sla-speech-tools.com>
- Muñoz, C. (2014). Contrasting effects of starting age and input on the oral performance of foreign language learners. *Applied Linguistics*, 35(4), 463–482. <https://doi.org/10.1093/applin/amu024>
- Muñoz, C. (2022). Audiovisual input in L2 learning. *Language, Interaction and Acquisition*, 13(1), 125–143. <https://doi.org/10.1075/lia.22001.mun>
- Nagle, C. L. (2023). A design framework for longitudinal individual difference research: Conceptual, methodological, and analytical considerations. *Research Methods in Applied Linguistics*, 2(1), 100033. <https://doi.org/10.1016/j.rmal.2022.100033>
- O'Brien, M. G., Derwing, T. M., Cucchiari, C., Hardison, D. M., Mixdorff, H., Thomson, R. I., Strik, H., Levis, J. M., Munro, M. J., Foote, J. A., & Levis, G. M. (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>
- Piske, T. (2007). Implications of James E. Flege's research for the foreign language classroom. In O.-S. Bohn, & M. J. Munro (Eds.), *Language experience in second language speech learning* (pp. 301–314). John Benjamins. <https://doi.org/10.1075/lllt.17.26pis>
- Rocca, B., Llompart, M., & Darcy, I. (2025). Phonological neighborhood density, phonetic categorization, and vocabulary size differentially affect the phonolexical encoding of easy and difficult L2 segmental contrasts. *Bilingualism: Language and Cognition*, 28(3), 662–675. <https://doi.org/10.1017/S1366728924000865>
- Rojczyk, A. (2015). Using FL accent imitation in L1 in foreign-language speech research. In E. Waniek-Klimczak, & M. Pawlak (Eds.), *Teaching and researching the pronunciation of English* (pp. 223–234). Springer. https://link.springer.com/chapter/10.1007/978-3-319-11092-9_14
- Saito, K. (2023). How does having a good ear promote successful second language speech acquisition in adulthood? Introducing auditory precision hypothesis-L2. *Language Teaching*, 56(4), 522–538. <https://doi.org/10.1017/S0261444822000453>
- Saito, K., Dewaele, J. M., & Abe, M. (2025). Disentangling the causal role of motivation, enjoyment, and anxiety in second language speech learning: A final report. *Studies in Second Language Acquisition*, 1–27. <https://doi.org/10.1017/S0272263125000038>
- Saito, K., Hanzawa, K., Petrova, K., Kachlicka, M., Suzukida, Y., & Tierney, A. (2022). Incidental and multimodal high variability phonetic training: Potential, limits, and future directions. *Language Learning*, 72(4), 1049–1091. <https://doi.org/10.1111/lang.12503>
- Saito, K., Kachlicka, M., Suzukida, Y., Mora-Plaza, I., Ruan, Y., & Tierney, A. (2024). Auditory processing as perceptual, cognitive, and motoric abilities underlying successful second language acquisition: Interaction model. *Journal of Experimental Psychology: Human Perception and Performance*, 50(1), 119–135. <https://doi.org/10.1037/xhp0001166>
- Saito, K., & Plonsky, L. (2019). Effects of second language pronunciation teaching revisited: A proposed measurement framework and meta-analysis. *Language Learning*, 69(3), 652–708. <https://doi.org/10.1111/lang.12345>
- Saito, K., Trofimovich, P., Abe, M., & In'nami, Y. (2020). Dunning-Kruger effect in second language speech learning: How does self-perception align with other perception over time? *Learning and Individual Differences*, 79, 101849. <https://doi.org/10.1016/j.lindif.2020.101849>

- 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/S0142716417000418>
- Sardegna, V. G. (2022). Evidence in favor of a strategy-based model for English pronunciation instruction. *Language Teaching*, 55(3), 363–378. <https://doi.org/10.1017/S0261444821000380>
- Scheffler, P., & Baranowska, K. (2023). Learning pronunciation through television series. *Language Learning and Technology*, 27(1), 1–16. <https://hdl.handle.net/10125/73520>
- Solon, M., Long, A. Y., & Gruzynski-Weiss, L. (2017). Task complexity, language-related episodes, and production of L2 Spanish vowels. *Studies in Second Language Acquisition*, 39(2), 347–380. <https://doi.org/10.1017/S0272263116000425>
- Stoughton, A. M., & Kang, O. (2024). A systematic review of empirical mobile-assisted pronunciation studies through a perception–production lens. *Languages*, 9(7), 251. <https://doi.org/10.3390/languages9070251>
- Strachan, L., Kennedy, S., & Trofimovich, P. (2019). Second language speakers' awareness of their own comprehensibility: Examining task repetition and self-assessment. *Journal of Second Language Pronunciation*, 5(3), 347–373. <https://doi.org/10.1075/jslp.18008.str>
- Szyska, M. (2022). Learners' attitudes to first, second and third languages pronunciation in structuring multilingual identity. *Applied Linguistics Review*, 13(6), 1127–1147. <https://doi.org/10.1515/applirev-2019-0061>
- 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., Isaacs, 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>
- Tsunemoto, A., Trofimovich, P., Blanchet, J., Bertrand, J., & Kennedy, S. (2022). Effects of benchmarking and peer-assessment on French learners' self-assessments of accentedness, comprehensibility, and fluency. *Foreign Language Annals*, 55(1), 135–154. <https://doi.org/10.1111/flan.12571>
- Tyler, M. D. (2019). PAM-L2 and phonological category acquisition in the foreign language classroom. In A. Nyvad, M. Hejná, A. Højen, A. Jespersen, & M. H. Sørensen (Eds.), *A sound approach to language matters: In honor of Ocke-Schwen Bohn* (pp. 607–630). Aarhus University Press. <https://ebooks.au.dk/aul/catalog/book/322>
- Uchihara, T., Karas, M., & Thomson, R. I. (2024). Does perceptual high variability phonetic training improve L2 speech production? A meta-analysis of perception-production connection. *Applied Psycholinguistics*, 45(4), 591–623. <https://doi.org/10.1017/S0142716424000195>
- Uchihara, T., Karas, M., & Thomson, R. I. (2025). High variability phonetic training (HVPT): A meta-analysis of L2 perceptual training studies. *Studies in Second Language Acquisition*, 1–34. <https://doi.org/10.1017/S0272263125100879>
- Wisniewska, N., & Mora, J. C. (2020). Can captioned video benefit second language pronunciation? *Studies in Second Language Acquisition*, 42(3), 599–624. <https://doi.org/10.1017/S0272263120000029>
- Xi, X., Li, P., & Prieto, P. (2024). Improving second language vowel production with hand gestures encoding visible articulation: Evidence from picture-naming and paragraph-reading tasks. *Language Learning*, 74(4), 884–916. <https://doi.org/10.1111/lang.12647>
- Xu, J., Saito, K., & Mora-Plaza, I. (2024). Task-based pronunciation teaching: Lack of auditory precision but not memory hinders learning. *System*, 127, 103532. <https://doi.org/10.1016/j.system.2024.103532>

Joan C. Mora is professor of English applied linguistics in the Department of Modern Languages and Literatures and English Studies at the University of Barcelona (UB) in Spain where he teaches courses in English phonetics and phonology and the acquisition of second language (L2) speech. His research aims at gaining a better understanding of how contextual and individual factors interact with pronunciation teaching methods and training techniques to shape L2 speech learning in instructed second language acquisition (SLA). His current research interests focus on the role of cognitive and emotional individual differences in the development of L2 pronunciation and speaking fluency, phonological learning in the mental lexicon, phonetic training methods, multimodal pronunciation training, and task-based pronunciation teaching and learning in instructed SLA.