Language Teaching Research Quarterly

EUROKD

2026 (55), 74–101
https://doi.org/10.32038/ltrq.2026.55.04

On English Language Learners' Prosody: The Possible Impact of L2 Aptitude, Musical Ability, and Proficiency on Nuclear Accent Placement in Spanish L1 Contexts

Christopher Felipe Barra–García
Faculty of Art and Humanities, Universidad de Concepción, Chile
Mauricio Véliz-Campos*
Faculty of Educational Sciences, Universidad de Talca, Linares, Chile
Mauricio A. Figueroa Candia
Faculty of Art and Humanities, Universidad de Concepción, Concepción, Chile
Yasna Pereira-Reyes
Faculty of Art and Humanities, Universidad de Concepción,Concepción, Chile
Mackarena Kartsevski Dune
Faculty of Educational Sciences, Universidad de Talca, Linares, Chile

Correspondence
Email: mauricio.veliz@utalca.cl

Abstract

Although the literature has recently recognised the significance of suprasegmental features in L2 English perception and production, nuclear accent (NA) placement remains partly terra incognita (Landblom & Ionin, 2022; Nagle & Baese-Berk, 2022). Thus, this study examined the extent to which Spanish L1 learners of English can perceive and produce English NA placement. A total of 53 participants completed NA placement perception (listening and identification) and production tasks based on utterances in unmarked and marked contexts. Also, the role of L2 aptitude, musical ability, and L2 proficiency was analysed using three standardised tests: the Modern Language Aptitude Test (MLAT), the Musical Ear Test (MET), and the Cambridge Placement Test. The results indicate that L2 aptitude was a weak predictor, whereas musical ability, specifically rhythm sensitivity, strongly predicted NA placement perception accuracy across the different sentence types. This effect was positively moderated by English proficiency. The study deepens our understanding of the perception-production link and supports the inclusion of prosodic training in L2 English acquisition.

ARTICLE HISTORY
Received: 19 April 2026
Revised: 22 July 2026
Accepted: 30 July 2026

KEYWORDS
Perception-Production, Nuclear Accent Placement, L2 Proficiency, L2 Aptitude, Musical Ability

How to cite this article (APA 7th Edition):
Barra–García, C. F., Véliz-Campos, M., Figueroa Candia, M. A, Pereira-Reyes, Y. , & Kartsevski Dune, M. (2026). On English language learners' prosody: The possible impact of L2 aptitude, musical ability, and proficiency on nuclear accent placement in Spanish L1 contexts. *Language Teaching Research Quarterly*, *55*, 74–101. https://doi.org/10.32038/ltrq.2026.55.04

Introduction

Second language (L2) pronunciation research has gone beyond isolated segmental contrasts to include suprasegmental phenomena including stress, rhythm, and intonation, given their significant role for intelligibility and communicative success (Calhoun et al., 2018; Landblom & Ionin, 2022; Mennen, 2015). Increasing attention has been paid to the perception-production relationship in L2 phonological acquisition, i.e., the extent to which learners' ability to perceive non-native contrasts relates to their ability to produce them accurately (Casillas, 2019; Flege, 1995, 2003; Melnik-Leroy et al., 2022; Shultz et al., 2012). Although this issue has been widely discussed in influential models such as PAM, PAM-L2, SLM, SLM-r, and L2LP, its developmental dynamics remain only partially understood, especially in the realm of suprasegments (Best, 1995; Best & Tyler, 2007; Flege, 1995; Flege & Bohn, 2021).

Suprasegmental features are indispensable to spoken communication because they shape the interpretation of information structure, discourse meaning, and speaker intent (Fadilla et al., 2023; Gordon & Darcy, 2018, 2022; Ladegaard & Chan, 2024; Yenkimaleki, 2019, 2021). In this respect, nuclearity plays a pivotal role in languages like English. Nuclear accent (NA) is often understood, following the British school of thought in particular, as the rightmost and most prominent pitch accent in an intonation group, capable of signalling focus and/or contrast (Halliday, 1967; Ladd, 1979; Levis & Wichmann, 2015). While both English and Spanish use prosodic prominence to signal information salience (or focus), they differ in terms of how they do so with English relying more heavily on prosody, whereas Spanish relies more heavily on syntax, and to a lesser extent, prosody. This makes nuclearity particularly challenging for Spanish speakers learning English (del Saz & Grau, 2022; Frota & Prieto, 2015; Hoot & Leal, 2020; Landblom & Ionin, 2022; Ortega-Llebaria & Colantoni, 2014; Zubizarreta & Nava, 2011).

In English prosody, the last lexical item (LLI) tends to attract the NA in unmarked utterances (del Saz & Grau, 2022; Halliday, 1967). Under certain circumstances, the NA is moved towards the left in an utterance, including cases of narrow focus, where part of the information in an utterance is treated as given information and de-accented as a result; also, in cases of broad focus, there are certain types of utterances where the NA is often found early on in the utterance, rather than on the LLI. Yet again, this poses additional challenges to Spanish speaking learners of English.

Spanish speaking learners' ability to perceive and produce NA may be partly explained by individual differences research (Bakkouche & Saito, 2026; Hosaka & Saito, 2026; Kidd et al., 2018; Li et al., 2022). In this study, the variables that have been included are L2 proficiency, L2 aptitude, and musical ability. L2 proficiency has repeatedly been associated with greater phonological sensitivity, better phonology-pragmatics alignment, and reduced L1 interference (Bachman & Palmer, 2010; Mennen, 2015; Ortega & Wu, 2025). It must be pointed out that L2 aptitude has largely been studied in relation to L2 learning success, yet its role in L2 pronunciation learning remains unexplored (Chalmers et al., 2021; Li, 2019; Saito, 2023). Musical ability, on the other hand, has drawn attention because of the shared auditory, temporal, and cognitive mechanisms involved in music and language processing (Patel, 2014; Swaminathan et al., 2021; Thompson et al., 2025).

Literature Review

The perception-production has attracted a good deal of attention in L2 phonological research. One of the longest-standing questions has been whether perception precedes production (Casillas, 2019, 2020; de Leeuw et al., 2021; Melnik-Leroy et al., 2022; Nagle, 2018). Major models including PAM, PAM-L2, SLM, SLM-r, and L2LP have provided influential accounts of how L1 categories shape L2 speech learning and how new phonetic categories may emerge over time (Best, 1995; Best & Tyler, 2007; Flege, 1995; Flege & Bohn, 2021). Much of the empirical basis for these models derives from segmental research, though, particularly on vowels, limiting their explanatory reach for suprasegmental acquisition (Gorba & Cebrian, 2021; Mennen & de Leeuw, 2014).

Although perception is often assumed to precede production, the empirical evidence is somewhat conflicting. Some studies show that perception gains can facilitate production gains, particularly in early development or following focused training (Casillas, 2019, 2020; Kachlicka, 2026; Melnik-Leroy et al., 2022; Nagle, 2018; Yenkimaleki, 2021). Others, however, suggest that learners may produce relatively target-like forms even when perceptual sensitivity remains limited, or that the two capacities develop asynchronously (de Leeuw et al., 2021; Kartushina et al., 2015; Kim & Han, 2024). The perception-production relationship is thus now widely seen as neither linear nor uniform across phonological domains, learners, or task conditions (Nagle & Baese-Berk, 2022), but appears to depend instead on the phonological target, task demands, prior experience, and other learner-internal variables.

One reason this issue remains unresolved is the comparatively limited attention given to suprasegmentals, even though suprasegmental features are central to how meaning is organised and interpreted in spoken language, contributing to intelligibility, discourse coherence, prominence marking, and pragmatic interpretation (Fadilla et al., 2023; Gordon & Darcy, 2018, 2022; Hosaka & Saito, 2026; Ladegaard & Chan, 2024; Nakabayashi, 2021). If the perception-production relationship is partly shaped by the complexity of the target feature, then prosodic phenomena provide a particularly important test case.

One such phenomenon is nuclear accent (NA) placement in English. NA is the most prominent accent in an intonational phrase and is closely tied to information structure, especially focus (Halliday, 1967; Ladd, 1979). As pointed out earlier, in both English and Spanish, prosody helps signal which element in an utterance is informationally salient: while broad-focus utterances typically all the information in an utterance as new, narrow-focus utterances place prominence more selectively on the leftward new element(s) (Ladd, 1979). NA placement is, therefore, both a phonetic and a discourse-pragmatic property, since it helps the listener interpret what is being foregrounded.

The canonical English pattern is for NA to fall on the last lexical item (LLI), often referred to as the default or unmarked pattern (Halliday, 1967; del Saz & Grau, 2022; Nilsenová & Swerts, 2012). This pattern covers many broad-focus utterances, especially when the final constituent is a content word such as a noun, adjective, adverb, or verb. English also allows marked cases in which NA shifts leftwards, though, leaving a de-accented tail to the right. These cases are especially relevant in L2 research because they require learners to move beyond a simple final-prominence strategy and process less transparent mappings between syntax, focus, and prosody. Existing work suggests that NA placement is a particularly difficult feature for Spanish-speakers learning English (Celce-Murcia et al., 2010; del Saz & Grau, 2022; Landblom & Ionin, 2022; Ortega-Llebaria & Colantoni, 2014; van Maastricht et al., 2016; Zubizarreta & Nava, 2011). Even so, research remains limited, particularly with regard to the joint examination of perception and production and to the learner variables that may support each.

The difficulty of English NA placement for Spanish-speaking learners is closely tied to cross-linguistic differences in focus realisation. Research on L1 influence has consistently shown that phonological learning is shaped by available structures and patterns available from the first language (Li et al., 2026). In Romance languages such as Spanish and French, syntactic flexibility often lets focused material shift into phrase-final position, aligning prosodic prominence with linear position more transparently (Frota & Prieto, 2015; Klassen, 2013; Landblom & Ionin, 2022; Zubizarreta & Nava, 2011). By contrast, Germanic languages such as English and German exhibit more rigid constituent order, which places greater pressure on prosody to encode in-situ focus (del Saz & Grau, 2022; Hirsch & Wagner, 2011; Hoot & Leal, 2020; Romero-Trillo, 2012). For Spanish-speaking learners, unmarked English utterances may not always pose serious difficulty because final prominence may still coincide with the target pattern. Marked utterances, however, require learners to override L1-based expectations and integrate discourse-pragmatic information in a more language-specific way. NA placement is thus especially well-suited to investigating both cross-linguistic influence and the developmental limits of L2 prosodic restructuring.

Although L1 influence is clearly important, it cannot fully explain variation in L2 prosodic performance. Individual differences also shape how learners process, acquire, and deploy phonological knowledge (Kidd et al., 2018; Lev-Ari, 2018; Li et al., 2022). L2 proficiency is one of the most robust and theoretically relevant of these variables. Proficiency is now widely understood as a multidimensional construct involving complexity, accuracy, and fluency (Li & Sui, 2025; Şahin Kizil, 2024; Skehan, 1998; Vercellotti, 2017). In phonological research, higher proficiency has been associated with more target-like production, greater perceptual sensitivity, and stronger integration of phonological and pragmatic information (Bachman & Palmer, 2010; Levis & Wichmann, 2015; Mennen, 2015; Ortega & Wu, 2025). Empirical studies have increasingly indicated the relevance of suprasegmentals to communication and comprehensibility (Gordon & Darcy, 2022; Saito, 2021; Saito & Plonsky, 2019; Sonsaat-Hegelheimer & Levis, 2025; Suzukida & Saito, 2022).

As far as L2 aptitude is concerned, it must be noted that it has been historically one of the most extensively studied predictors of language learning success, together with working memory, and the MLAT remains one of its best-established measures (Carroll & Sapon, 2002; Hosaka & Saito, 2026; Riquelme et al., 2025; Skehan & Chen, 2024; Sparks, 2022; Sparks & Dale, 2023a; Sparks et al., 2023; Wen et al., 2017). The MLAT has demonstrated strong reliability and validity and has been widely used for studying grammar, vocabulary learning, and overall achievement (Chalmers et al., 2021; Dale & Sparks, 2024; Li, 2019; Sparks & Dale, 2023b). Nonetheless, its role in pronunciation learning remains rather unclear. Recent empirical research indicates that linking aptitude to L2 pronunciation is still relatively sparse and methodologically heterogeneous (Saito, 2023; Trofimovich et al., 2015).

Musical ability has emerged as another promising variable in L2 phonological research because of the shared mechanisms underlying musical and linguistic processing. Work in this area has pointed to links between musicality and pitch perception, temporal processing, auditory attention, and memory, all of which are relevant to speech learning (Bidelman et al., 2013; Ghooch Kanloo et al., 2026; Patel, 2014; Turker & Reiterer, 2021). In L2 contexts, musical ability has been associated with enhanced perception and, in some cases, improved production of non-native contrasts, especially where timing, pitch, or prosodic patterning is involved (Cason et al., 2020; Delogu & Zheng, 2020; Götz et al., 2023; Li & DeKeyser, 2017; Milovanov et al., 2010). The Musical Ear Test (MET) operationalises musical ability through melody and rhythm (Wallentin et al., 2010), a highly relevant distinction for NA placement. Melody may relate to pitch-based prominence, whereas rhythm may be associated with temporal structuring and speech timing. Because Spanish and English differ rhythmically, and because NA placement is embedded within broader prosodic organisation, rhythm sensitivity may be particularly relevant to learners' ability to detect prominence in English. However, the strength and nature of the relationship between musicality and L2 phonology remain unsettled, especially for suprasegmental features and for the distinction between perception and

production (Chobert & Besson, 2013; Patel, 2014; Saito et al., 2026; Thompson et al., 2025).
The present study addresses these gaps by examining how Spanish L1 learners of English perceive and produce NA placement across different sentence types and by exploring the extent to which this performance is associated with L2 proficiency, L2 aptitude, and musical ability.

Methodology

Research Design

This study employed a quantitative, non-experimental, cross-sectional design with descriptive, comparative, correlational, and predictive components. It was non-experimental because no variable was manipulated and no participants were assigned to treatment conditions. It was cross-sectional because all data were collected within a single phase of measurement. The study was descriptive in characterising learners' performance in the perception and production of English nuclear accent (NA) placement, comparative in examining differences between perceptual and productive performance across unmarked and marked sentence types, and correlational-predictive in investigating the extent to which L2 proficiency, L2 aptitude, and musical ability were associated with and predicted NA performance. The design therefore aimed to identify patterns of association among naturally occurring variables rather than to establish causal relationships.

Research Questions

The study addressed the following research questions:

RQ$_1$: To what extent are Spanish L1 learners of English able to perceive and produce English nuclear accent placement?
RQ$_2$: What is the relationship between learners' perception and production of English nuclear accent placement and L2 aptitude?
RQ$_3$: What is the relationship between learners' perception and production of English nuclear accent placement and musical ability?
RQ$_4$: To what extent does L2 proficiency relate to learners' perception and production of English nuclear accent placement?

Participants

Participants in this study were 53 Chilean undergraduate students enrolled in an English teacher education programme. A small subgroup (n = 9) attended a different campus but followed the same curriculum. All participants were either currently taking, or had already completed, a university-level intonation course that included theoretical and practical work on English NA placement.

The sample was obtained through non-probability convenience sampling from an accessible student population. To ensure internal consistency, only participants who met the following criteria were included: (a) Spanish as their first language, (b) no selfreported sensory, hearing or cognitive impairment likely to affect task performance, and (c) completion of all five instruments administered in the study. English proficiency was also measured in order to characterise the sample and include proficiency as an explanatory variable in the analyses. All participants took part on a voluntary basis, signed a consent form and were informed of their right to withdraw at any stage of the study without penalty.

Instruments

Five computer-based instruments were administered in this study: a Cambridge online general English proficiency test, two researcher-designed nuclear accent placement tasks, the first two tasks of the Modern Language Aptitude Test (MLAT), and the Musical Ear Test (MET). Together, these instruments operationalised the dependent variables of the study, namely English nuclear accent (NA) placement perception and production, and the three learner-related variables examined as potential predictors or covariates, viz. L2 proficiency, L2 aptitude, and musical ability.

English proficiency measure

To measure participants' English proficiency, the Cambridge online General English Test was applied. This test belongs to the Cambridge English assessment battery and is designed as a brief diagnostic instrument that helps learners identify their approximate level of English and the Cambridge English examination pathway most suited to them. In the present study, the test was used as a practical external measure of general English proficiency that could be incorporated into the statistical models as an explanatory variable.

The decision to include a proficiency measure was theoretically and methodologically necessary. First, L2 proficiency has been repeatedly associated with phonological sensitivity, target-like production, and the capacity to integrate linguistic and pragmatic information in L2 speech. Second, while the study involved learners from the same teacher education programme, they did not necessarily have the same English levels. Measuring proficiency thus allowed the study to characterise the sample and to examine whether variation in NA placement performance could be – at least in part - explained by participants' general English development rather than by – or in conjunction with - language aptitude or musical ability alone.

The test was administered online in computer-based format. It consisted of 25 compulsory multiple-choice items including grammar, vocabulary, and collocation items. Each correct answer was awarded one point, yielding a maximum score of 25. The first five items presented three response options, while the remaining 20 items presented four response options; answers required the selection of one to three words. Once the test was completed, participants obtained an overall score on the screen with an interpretation of results with reference to CEFR levels. In the present sample, scores were used

quantitatively as a continuous proficiency index and descriptively to characterise participants' overall English level.
The test was selected because it provided a brief, standardised, externally developed, and easily administrable measure of general English proficiency. It was also suitable for group administration under computer-laboratory conditions and placed minimal additional burden on participants. However, its scope should be interpreted with caution. The instrument does not provide a full four-skills proficiency profile, nor does it include extended speaking, writing, or reading-comprehension tasks. For this reason, it was treated as a general proficiency covariate rather than as a comprehensive measure of communicative competence. In terms of validity, its relevance for the present study rests on its alignment with general English proficiency descriptors and its usefulness as an external indicator of participants' global language level. As no item-level reliability statistics were available for the public online version of the test, no internal reliability coefficient is reported for this instrument in the present study.

Nuclear accent perception and production tasks

Two researcher-designed instruments were used to assess participants' ability to perceive and produce English NA placement. These tasks were developed specifically for the present study, as no readily available standardised instrument was found that directly assessed English NA placement in Spanish L1 learners across both unmarked and marked broad-focus utterances. Their design was grounded in the theoretical literature on English sentence stress, information structure, broad focus, and the Last Lexical Item (LLI) rule, as well as in previous research showing that English NA placement is especially challenging for Spanish-speaking learners due to cross-linguistic differences in the prosodic and syntactic realisation of focus.

Both tasks contained the same set of 35 broad-focus utterances. Each utterance contained between five to seven words. Of these, 19 utterances followed the LLI rule and were therefore classified as unmarked. The remaining 16 utterances represented marked cases or exceptions to the LLI pattern. These included event sentences, wh-questions ending in verbs, utterances involving final function words or empty words, final adverbials of time and place, reflexive pronouns, and final vocatives. Including both unmarked and marked sentence types was key to the construct being measured. Unmarked utterances allowed the study to assess learners' ability to identify or produce the canonical English tendency for NA to fall on the final lexical item, while marked utterances assessed whether learners could move beyond a simple final-prominence strategy and assign prominence to an earlier constituent as required by the discourse-prosodic structure of the utterance.

Examples of items:
a) Audio for LLI rule (unmarked): I believe they will celebrate their <u>**birthday**</u>
b) Audio for exceptions (marked): They signed an <u>**agreement**</u> last year.

The sentence set was developed and reviewed by three researchers with expertise in English phonetics, phonology, and L2 speech. This expert review strengthened content validity by ensuring that the items represented the intended NA categories and that the marked items genuinely required a departure from the default LLI pattern. The two tasks were also matched in number of items, sentence length, and sentence type, improving comparability between perception and production instruments. All stimuli were randomised for each presentation of the task. Audio stimuli were recorded in mono .WAV format at a sampling rate of 44.1 kHz and 16-bit depth.

Several steps were taken to strengthen reliability and validity. Content validity was addressed via expert construction and checking of the sentence set. Construct validity was addressed by designing the tasks around theoretically motivated NA categories: unmarked LLI items and marked exceptions requiring leftward displacement of nuclear prominence. Procedural reliability was supported by computer-based administration, automatic item progression, fixed exposure time, randomised presentation, and standardised audio specifications. For the perception task, scoring reliability was supported by automatic recording of dichotomous responses by the platform. For the production task, a total of 1,855 recorded sentences were independently assessed by two experienced teachers of English as an L2 who used Praat software to double-check parameters for NA placement (mainly, Fundamental frequency). Inter-rater reliability was evaluated using a two-way random-effects intraclass correlation coefficient with absolute agreement, ICC (2,1). The analysis showed excellent agreement between raters, ICC (2,1) = .962, 95% CI [.935, .978]. The mean difference between raters was not statistically significant, M_{diff} = 0.019, $t(52) = 0.17$, $p = .868$, indicating no evidence of systematic rater bias. Bland–Altman limits of agreement ranged from -1.589 to 1.626, suggesting minimal disagreement across the scoring range. These results indicate a high degree of reliability for the two raters' production measures.

L2 aptitude measure

To assess participants' L2 aptitude, this study used two sections of the Modern Language Aptitude Test (MLAT; Carroll & Sapon, 2002). The MLAT was originally developed by John B. Carroll and Stanley Sapon in the 1950s as part of their research on predicting foreign language learning success, and has been widely used to measure language aptitude in L2 acquisition research. The test assumes the premise that language aptitude is not an isolated general ability, but rather a composite of interconnected variables, including phonetic coding ability, grammatical sensitivity, rote learning ability, and inductive language learning ability.

The full MLAT consists of five parts altogether: Number Learning, Phonetic Script, Spelling Clues, Words in Sentences, and Paired Associates. These subtests tap into different abilities often associated with foreign language learning. However, the present study administered only Part I, Number Learning, and Part II, Phonetic Script. This decision was made for reasons of construct relevance. The remaining sections could also confound the construct, as they may rely on language proficiency. As the current study focused on L2 prosodic perception and production, the aim was to prioritise aptitude components more directly related to auditory processing, memory for novel sound-meaning associations, and sound-symbol learning.

Part I, Number Learning, measures individuals' ability to learn and remember number words in an unfamiliar language-like system through auditory stimulus. This part addresses the use of auditory memory, sound-meaning association, and the ability to retain and manipulate unfamiliar verbal material. The maximum score in Part I is 15. Part II, Phonetic Script, measures the ability to associate speech sounds with unfamiliar written phonetic symbols. It is composed of three sets, each with five subsets of four nonsense syllables. The symbols used are not the International Phonetic Alphabet, although some may resemble familiar phonetic notation. Part II has a maximum score of 29. The maximum combined score for the two sections used in this study was therefore 44. Each correct answer was awarded one point, and no deductions were made for errors or omissions. Scores were calculated according to the official MLAT answer key.

The MLAT was selected because of its long history, extensive use in L2 research, and strong reputation as a foreign language achievement predictor. Previous work has consistently treated the MLAT as one of the best-established measures of language aptitude, and subsequent research has continued to support its relevance for explaining differences in L2 learning outcomes. At the same time, because only Parts I and II were used in the present study, the aptitude construct operationalised here should be interpreted as sound-related and memory-related L2 aptitude rather than as global language aptitude. This narrower operationalisation was appropriate for our research purposes because English NA placement requires learners to process auditory prominence patterns, retain unfamiliar prosodic mappings, and coordinate sound-related information with linguistic structure. The use of selected MLAT sections thus increased construct relevance while reducing overlap with general English proficiency.
Concerning reliability and validity, the MLAT has been widely reported to have strong psychometric properties and has been used extensively in studies of L2 achievement, grammar, vocabulary, and general language learning success. Its predictive validity has been supported by research showing associations between MLAT scores and L2 course performance or later L2 outcomes. In the present study, the use of only two sections means that reliability and validity claims must be interpreted at the selected components' level rather than the full MLAT battery. The decision to use these sections was therefore both theoretically motivated and methodologically conservative.

Musical ability measure

Musical ability was assessed using the Musical Ear Test (MET; Wallentin et al., 2010). The MET has been described as a short, objective, and accessible measure of musical competence for musicians and non-musicians. It is administered relatively quickly and allows to differentiate between individuals with different levels of musical ability. The test focuses on perceptual discrimination and has shown to be efficient to assess individuals with wide variations in musical training levels.

The MET was selected for the present study because it provides separate measures of melody and rhythm sensitivity, both of which are theoretically relevant to L2 prosody. Melody sensitivity is associated with pitch and contour perception, which may contribute to detecting prominence and intonational movement. Rhythm sensitivity is associated with temporal structuring and auditory patterning, which are also central to speech rhythm and prosodic organisation. As English and Spanish differ in their rhythmic and prosodic organisation, and since NA placement depends on learners detecting prominence within the speech stream, the MET provided an appropriate way of examining whether individual differences in musical perception were associated with NA placement perception and production.

The MET contains two subtests: melody and rhythm. Each subtest presents 52 pairs of stimuli, giving a total of 104 scored items. In each trial, two short musical sequences are presented. Participants have to decide whether they are the same or different. Each correct answer is awarded one point. Half of the trials in each subtest are identical and half are different, and item order is randomised. Each subtest begins with two unscored practice trials with feedback, after which no feedback is provided during the test itself. The test lasts around 20 minutes and is played continuously. Participants register their answers on an answer sheet by crossing out or shading a slot under the *Same* or *Different* column.

The melody subtest uses short piano melodies to assess pitch and contour discrimination. Each melody lasts one measure, contains three to eight tones, and is played at 100 beats per minute. In the different trials, one pitch is altered; in some cases, this pitch violation also produces a contour violation. The melody set includes non-diatonic tones as well as melodies in major and minor keys. The rhythm subtest presents short rhythmic sequences played with wood-block sounds. Each rhythmic sequence contains four to 11 beats, lasts one measure, and is also played at 100 beats per minute. These design features allow the MET to capture individual differences in both melodic and rhythmic discrimination while avoiding dependence on specialised musical notation or music-theory terminology. All the recordings used in the current study were provided by Dr. Mikkel Wallentin and Dr. Peter Vuust. Vuust kindly helped with various questions on the test details and procedures.

The MET has shown strong evidence of reliability and validity. In the original validation study, the full test showed high internal consistency, with Cronbach's α reported at .87. The melody and rhythm subtests also showed high internal consistency, with α values of .96 and .94, respectively. The test was able to distinguish between non-musicians, amateur musicians, and professional musicians, and MET scores were associated with musical practice and auditory memory measures. Later work with larger samples has also supported its use as an objective measure of musical ability, reporting approximately normal score distributions and limited floor or ceiling effects. These properties made the MET particularly suitable for the present study, whose sample was not selected on the basis of musical training and whose aim was to examine naturally occurring variation in musical perception.

Procedure

Data collection took place in university computer laboratories under controlled conditions. Participants completed the tasks individually using desktop or laptop computers equipped with headsets and microphones. Testing conditions were standardised as far as possible across sessions, and participants were allowed to adjust audio volume as needed. Each session was supervised by two invigilators. Data collection was distributed across several days to reduce fatigue, but the order of administration remained constant for all participants. When two instruments were completed in the same session, participants were given a short break of approximately 5 to 10 minutes.

English proficiency measure

The General English Test was presented in an online computer-based format in a computer laboratory. Participants were instructed to answer the 25 items in the test and requested not to close the final screen of their results so that the invigilator could register their results. There is no time restriction suggested for this test; however, participants were advised to read the items carefully and to answer at their own pace. This was the first test given to participants.

Nuclear accent perception and production tasks

The production task was administered first, using a computer. The test was created and programmed on a devoted platform. This decision was made to reduce the possibility that listening to the target utterances in the perception task would prime subsequent oral production. In each production trial, sentences were presented individually on screen with a 30-second window in which to rehearse and produce it orally. During this window, the platform automatically recorded the response from the participant. Participants could rehearse aloud, but only the final complete reading of each item was retained for further analysis. Once the time limit elapsed, the written item disappeared from the screen and the platform automatically advanced to the next sentence. Recorded responses were scored dichotomously: a fully accurate NA placement received one point, while inaccurate, incomplete, unclear, or partial responses received zero points. The maximum production score was therefore 35, with separate possible sub-scores of 19 for unmarked utterances and 16 for marked utterances. A total of 1,855 recorded sentences were obtained.

The perception task was administered after the production task, using a computer and headset. This test was also created and programmed on a devoted platform. Its purpose was to measure participants' ability to listen to and identify the location of NA in 35 recorded English utterances. Each item was played only once with no possibility of replaying the stimulus, although participants could read the sentence on screen while selecting their answer from provided written options. Each trial presented six response options, corresponding to three words from the utterance, each paired with a tonal alternative (Rise or Fall). Participants had 30 seconds to select and submit their response. If a participant clicked an option but did not formally submit it before the end of the time window, the platform still recorded the selected alternative. If no option was clicked, the item was scored as incorrect. As in the production task, responses were scored dichotomously, with one point for a correct answer and zero points for an incorrect answer. The maximum perception score was 35, allowing separate scores for the marked and unmarked patterns. Participants were not informed of their results once the test was completed.

L2 aptitude measure

Part I and Part II of the MLAT test were presented to participants using the original test audios and the written questions. Part I (Number learning) used 15 recorded items and Part 2 (Phonetic script) had 29 recorded items. Each part of the test is preceded by examples to present sample audio material and train participants on how to answer the test questions. All answers are registered on paper. The total maximum test score is 44 points. The presentation of items, audios, answer sheets and scoring procedure followed the standard MLAT guidelines. Participants were not informed on their results.

Musical ability measure

The MET test was given to participants using the original audios for melody and rhythm items. Each component was comprised of 52 audio stimuli. The test started with the melody component; instructions were presented in audio in Spanish. Two examples (one stimulus with same 2 melody items and one with two different melody items) were then presented before the test items. Participants recorded their answers (same or different) by crossing out a slot under the corresponding column. The rhythm part was presented after the melody test items. The same procedure was performed. The audio of the test was played non-stop for both MET components. Participants were not informed on their final score.

Data Analysis

All data were analysed quantitatively. Cambridge placement test scores were retrieved from the official online platform. Perception responses were automatically recorded as correct or incorrect, and MLAT and MET scores were calculated using the relevant scoring keys. Production data were independently rated by two experienced teachers of English, and inter-rater reliability was assessed using a two-way random-effects intraclass correlation coefficient with absolute agreement (ICC[2,1]).

The analysis proceeded in four stages. First, descriptive statistics were calculated for all main variables. Secondly, paired-samples *t* tests compared perception and production scores across overall, unmarked, and marked conditions. Thirdly, Pearson correlations examined zero-order associations among L2 proficiency, L2 aptitude, musical ability, and NA perception and production. Finally, multiple linear regression analyses were conducted to assess the predictive contribution of L2 proficiency, L2 aptitude, and musical ability to learners' NA perception and production scores, including interaction terms where appropriate. Variables were standardised as z-scores for the regression analyses.

Ethical Considerations

The study received formal approval from the relevant institutional bioethics committee. All procedures adhered to standard ethical principles for research involving human participants, including informed consent, voluntary participation, confidentiality, and the right to withdraw without consequence.

Results

Descriptive Results

Table 1

Descriptive Statistics for L2 Proficiency, L2 Aptitude, Musical Ability, and Nuclear Accent Perception and Production

Variable	*N*	*M*	% Correct	*SD*	Min	Max	*W*	*p*
L2 Proficiency	53	18.30	73	2.89	8	23	0.940	< .05
Overall Musical Ability	53	69.00	66	9.07	45	88	0.984	.714
Melody Sensitivity	53	33.30	64	4.84	21	45	0.977	.412
Rhythm Sensitivity	53	35.70	68	5.68	24	46	0.971	.222
Overall L2 Aptitude	53	31.90	72	5.54	21	41	0.954	< .05
Number Learning	53	7.43	49	4.66	0	15	0.917	< .01
Phonetic Script	53	24.40	84	2.24	19	28	0.946	< .05
Overall NA Perception	53	15.20	43	5.36	5	27	0.980	.505
Perception (Unmarked)	53	11.40	59	4.04	3	18	0.967	.142
Perception (Marked)	53	3.81	23	2.72	0	11	0.912	< .001
Overall NA Production	53	19.30	55	2.97	13	25	0.975	.320
Production (Unmarked)	53	14.30	75	2.24	9	18	0.954	< .05
Production (Marked)	53	5.02	31	1.97	2	11	0.925	< .01

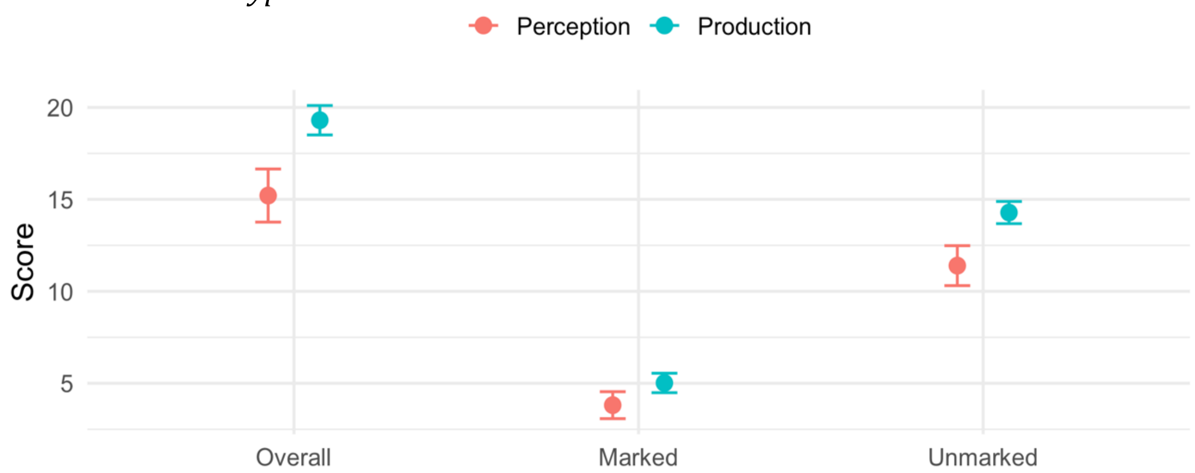
Note. NA = nuclear accent. Percentages are included to facilitate interpretation across measures with different score ranges. Maximum possible scores were as follows: Cambridge Placement Test = 25; overall musical ability (MET total) = 104; melody sensitivity = 52; rhythm sensitivity = 52; overall L2 aptitude = 43; number *learning = 15; phonetic script = 28; overall NA perception = 35; perception (unmarked) = 19; perception (marked) = 16; overall NA production = 35; production (unmarked) = 19; production (marked) = 16.*

Perception–production comparisons

A paired-samples t test was conducted to compare NA placement perception and production across overall, unmarked, and marked conditions. Across all three comparisons, production scores were significantly higher than perception scores. Overall, production exceeded perception, $t(52) = 5.07$, $p < .001$, $d = 0.70$, 95% CI [0.39, 0.99]. The same pattern was observed in unmarked contexts, $t(52) = 4.50$, $p < .001$, $d = 0.62$, 95% CI [0.32, 0.91], and in marked contexts, although with a smaller effect size, $t(52) = 2.75$, $p < .01$, $d = 0.38$, 95% CI [0.10, 0.66].

Figure 1

Mean NA Placement Perception and Production Accuracy Across Overall, Unmarked, and Marked Sentence Types



Correlations among the main variables

Table 2 presents the zero-order correlations among L2 proficiency, overall musical ability, overall L2 aptitude, and the perception and production measures. At the bivariate level, overall NA perception was positively associated with L2 proficiency ($r = .36$, $p < .01$), overall musical ability ($r = .37$, $p < .01$), and overall L2 aptitude ($r = .31$, $p < .05$). By contrast, overall NA production was not significantly correlated with any of these variables. A similar pattern emerged across sentence types: perception in marked contexts was positively associated with overall musical ability ($r = .34$, $p < .05$), whereas perception in unmarked contexts was positively associated with L2 proficiency ($r = .37$, $p < .01$) and overall L2 aptitude ($r = .35$, $p < .05$). No corresponding significant associations were observed for production in either marked or unmarked contexts.

Table 2

Correlations among the Main Variables

Variable	1	2	3	4	5	6	7	8	9
1. L2 Proficiency	—								
2. Overall Musical Ability	.14	—							
3. Overall L2 Aptitude	.37**	.28*	—						
4. Perception (Overall)	.36**	.37**	.31*	—					
5. Production (Overall)	.06	−.18	−.16	.09	—				
6. Perception (Marked)	.16	.34*	.10	.68***	.13	—			
7. Production (Marked)	.21	−.23	.05	.12	.66***	.10	—		
8. Perception (Unmarked)	.37**	.26	.35*	.87***	.04	.23	.09	—	
9. Production (Unmarked)	−.10	−.04	−.25	.02	.75***	.08	−.01	−.03	—

Note. p < .05, p < .01**, p < .001***.*

Predictors of NA perception

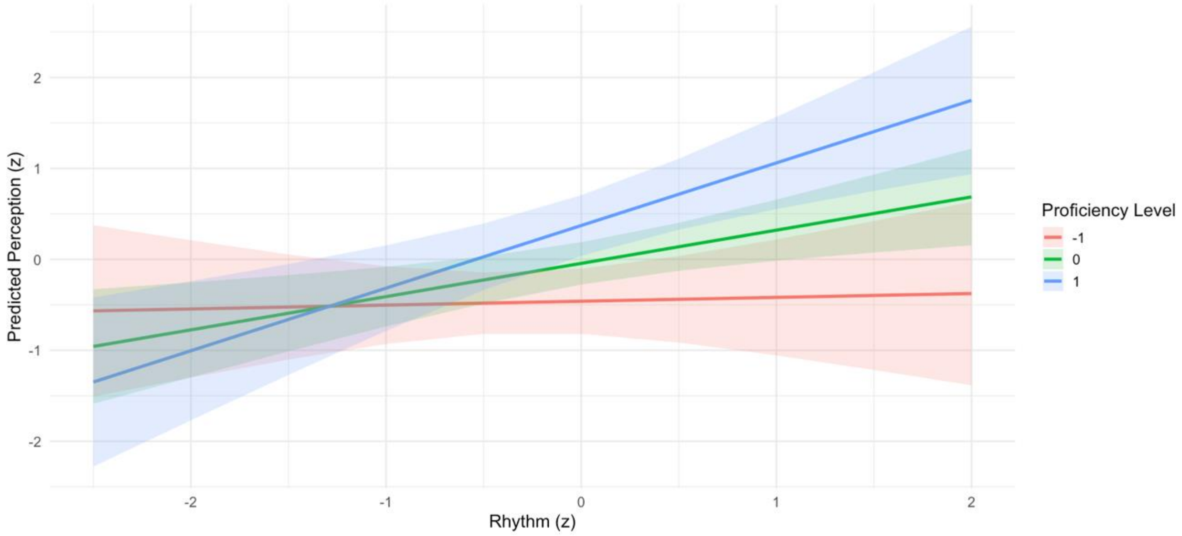
A series of multiple linear regression analyses was conducted to examine whether L2 proficiency, L2 aptitude, and musical ability predicted NA perception. In the overall perception model, the set of predictors was significant, $F(3, 49) = 5.30$, $p < .01$, accounting for 20% of the variance. Within this model, L2 proficiency ($\beta = 0.27$, $t = 2.04$, $p < .05$) and overall musical ability ($\beta = 0.29$, $t = 2.25$, $p < .05$) were significant positive predictors, whereas overall L2 aptitude was not ($\beta = 0.13$, $t = 0.93$, $p = .357$).

A second model tested the interaction between L2 proficiency and overall musical ability. Although the model was significant, $F(3, 49) = 5.72$, $p < .01$, explaining 21% of the variance, the interaction term did not reach significance ($\beta = 0.21$, $t = 1.35$, $p = .182$). When the musical ability composite was disaggregated, only rhythm sensitivity significantly predicted overall NA perception, $F(2, 50) = 6.19$, $p < .01$, $R^2 = .17$; rhythm sensitivity: $\beta = 0.47$, $t = 3.26$, $p < .01$; melody sensitivity: $\beta = -0.06$, $t = -0.42$, $p = .677$. A final model showed that the interaction between rhythm sensitivity and L2 proficiency was significant, $F(3, 49) = 8.75$, $p < .001$, $R^2 = .31$; $\beta = 0.32$, $t = 2.15$, $p < .05$, indicating that the positive effect of rhythm sensitivity on overall NA perception increased as proficiency increased (Figure 2).

The analyses by sentence type revealed a similar but more differentiated pattern. In unmarked contexts, the initial model was significant, $F(3, 49) = 4.46$, $p < .01$, $R^2 = .17$, with L2 proficiency emerging as the only significant predictor ($\beta = 0.28$, $t = 2.03$, $p < .05$). When melody and rhythm were examined separately, only rhythm sensitivity remained significant, $F(2, 50) = 3.68$, $p < .05$, $R^2 = .09$; $\beta = 0.40$, $t = 2.64$, $p < .05$. The interaction between rhythm sensitivity and proficiency was also significant in this context, $F(3, 49) = 8.02$, $p < .001$, $R^2 = .29$; $\beta = 0.42$, $t = 2.74$, $p < .01$ (Figure 3).

Figure 2

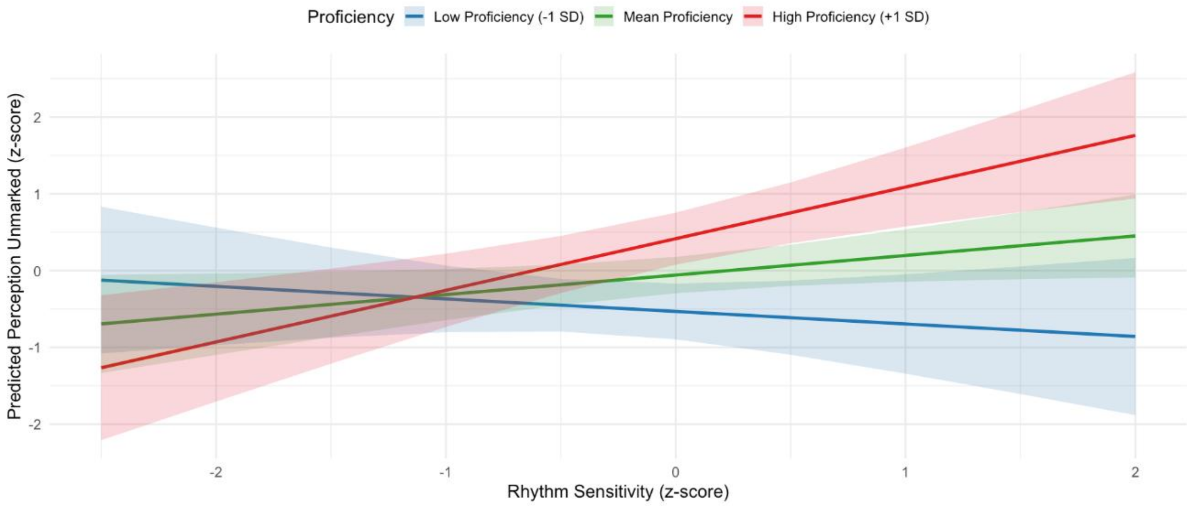
Interaction Effect of Rhythm Sensitivity and L2 Proficiency on Overall NA Placement Perception Accuracy



Note. Error bars show 95% confidence intervals.

Figure 3

Interaction Between Rhythm Sensitivity and L2 Proficiency Predicting NA Placement Perception Accuracy in Unmarked Contexts



Note. Error bars represent 95% confidence intervals.

In marked contexts, the initial model was only marginally significant, $F(3, 49) = 2.37$, $p = .082$, $R^2 = .07$. In this model, overall musical ability predicted marked-context perception ($\beta = 0.33$, $t = 2.36$, $p < .05$), whereas L2 proficiency and overall L2 aptitude did not. When musical ability was disaggregated, only rhythm sensitivity remained significant, $F(2, 50) = 3.75$, $p < .05$, $R^2 = .10$; $\beta = 0.34$, $t = 2.24$, $p < .05$. Unlike the overall and unmarked models, however, the interaction between rhythm sensitivity and L2 proficiency was not significant in marked contexts ($\beta = 0.02$, $t = 0.10$, $p = .924$).

Taken together, the perception analyses indicate that rhythm sensitivity was the most consistent predictor of NA perception, and that its contribution was amplified by higher L2 proficiency, particularly in overall and unmarked conditions.

Predictors of NA production

Inter-rater reliability for the production ratings was excellent, ICC(2,1) = .962, 95% CI [.935, .978]. The mean difference between raters was not significant, $t(52) = 0.17$, $p = .868$, and Bland–Altman limits of agreement ranged from −1.589 to 1.626, indicating minimal disagreement.

In contrast to the perception analyses, none of the regression models for NA production reached significance. The initial overall production model was non-significant, $F(3, 49) = 1.09$, $p = .362$, $R^2 = .01$, and none of the predictors contributed significantly: L2 proficiency ($\beta = 0.14$, $t = 0.97$, $p = .335$), overall musical ability ($\beta = -0.15$, $t = -1.07$, $p = .290$), and overall L2 aptitude ($\beta = -0.17$, $t = -1.08$, $p = .283$).

The same pattern held when production was analysed by sentence type. In unmarked contexts, the model was non-significant, $F(3, 49) = 1.13$, $p = .344$, $R^2 = .01$, although L2 aptitude showed a marginal tendency ($\beta = -0.26$, $t = -1.68$, $p = .099$). In marked contexts, the model was also non-significant, $F(3, 49) = 2.05$, $p = .119$, $R^2 = .06$, with only a marginal effect for overall musical ability ($\beta = -0.27$, $t = -1.93$, $p = .060$).

Finally, two additional models were used to test whether perception predicted production and whether production predicted perception. Neither model was significant, $F(1, 51) = 0.45$, $p = .504$, indicating no evidence of a bidirectional predictive relationship between the two domains in the present dataset.

Discussion

This study investigated how Chilean Spanish L1 learners of English perceive and produce English nuclear accent (NA) placement, and whether their performance is associated with L2 aptitude, musical ability, or L2 proficiency. Three major findings emerged: (i) English NA placement remains only partially mastered by these learners, particularly in marked contexts - that is, where NA does not fall on the last lexical item; (ii) perception and production seem to operate differently, functioning not as two symmetrical manifestations of a single underlying ability; and (iii) learner-internal variables do not contribute in equal measure: rhythm sensitivity and proficiency appear to better predict perception, whereas MLAT aptitude measures showed very little predictive power. These findings thus suggest that NA placement is best understood not merely as a pronunciation feature but as an interface phenomenon involving prosody, syntax, information structure, auditory processing, and L2 developmental experience - an interpretation which corroborates recent work indicating that L2 speech learning is shaped not only by cross-linguistic influence but also by perceptual, cognitive, auditory-processing, and experiential differences among learners (Bakkouche & Saito, 2026; Hosaka & Saito, 2026).

Why marked NA placement proved particularly difficult needs to be further unpacked. NA placement, within the conceptual framework adopted here, is a prosodic mechanism through which speakers structure information and signal what is perceived as or meant to be communicatively salient (Halliday, 1967; Ladd, 1979). This process seems highly influenced by learners' L1 sentence-accent patterns and this factor may contribute to clarify which contexts pose greater challenge for perception and production of NA. In unmarked utterances in English, there is a tendency to place NA on the last lexical item. This pattern can coincide with a general prominence pattern and with learners' L1 sentence-accent expectations, which can make the target pattern relatively transparent. Cross-linguistic influence quite possibly exacerbates the difficulty of NA placement for L2 speakers and listeners. Spanish and English both draw on prosodic prominence to express informational salience, yet they differ in the way syntax and prosody interact to achieve communicative salience. Spanish favours LLI NA placement even in narrow-focus contexts (del Saz & Grau, 2022) and allows greater syntactic flexibility, often allowing focused components to surface in utterance-final position where prosodic prominence aligns with linear order. English exhibits a less flexible word order and relies more heavily on prosody to mark focus (Frota & Prieto, 2015; Zubizarreta & Nava, 2011). The learners' challenge, then, is best read as one of restructuring the mapping between form, position, and discourse function, instead of a lack of capacity to hear or produce NA. This interpretation is consistent with previous work showing that Spanish-speaking learners often struggle with English accentual patterns that do not follow the LLI rule (del Saz & Grau, 2022; Landblom & Ionin, 2022; Ortega-Llebaria & Colantoni, 2014).

The second major finding - that production exceeded perception - calls for cautious interpretation. On its face, this result appears to challenge perception-first accounts of L2 speech development, particularly models in which perceptual category formation is expected to underpin subsequent production (Best & Tyler, 2007; Flege & Bohn, 2021). It should not, however, be interpreted as evidence that production is generally more developed than perception in L2 prosody. A more plausible interpretation is that the two tasks placed different demands on learners altogether. The perception task required participants to identify NA placement after a single auditory presentation, thus placing a heavy burden on online auditory discrimination, attention to prosodic cues, and immediate mapping between acoustic prominence and information structure. In contrast, the production task allowed for rehearsal and may have let learners draw on explicit knowledge acquired in their intonation course. Higher production scores may thus reflect more controlled, rule-based performance under favourable task conditions rather than fully automatised productive command of English NA placement. This reading reinforces the broader argument that the perception-production relationship is neither linear nor uniform, nor necessarily bidirectional, across all phonological domains

(Nagle, 2018; Nagle & Baese-Berk, 2022): perception and production in NA placement may rely on partially distinct mechanisms, since perception requires the listener to detect and interpret prominence in real time, while production requires the speaker to plan and execute a prominence pattern in relation to syntax and discourse.

The limited predictive power of L2 aptitude also calls for a critical discussion. The absence of significant effects for MLAT-based aptitude should not be taken as evidence that cognitive abilities are irrelevant to L2 prosody altogether; more likely, the result points to a measurement and construct-validity issue, in that the specific aptitude components assessed may not fully capture the cognitive mechanisms most relevant to NA placement. The MLAT has traditionally proven useful in predicting broader language-learning outcomes, including sound-symbol association, memory, and analytic language-learning ability (Carroll & Sapon, 2002; Wen et al., 2017). NA placement, however, taps into sensitivity to temporal patterning, prominence distribution, cue integration, and the interface between phonology and discourse. These demands may be closer to auditory-temporal processing and prosodic attentional control than to the abilities measured by the MLAT subtests used here, as Hosaka and Saito (2026) have suggested.

Rhythm sensitivity, by contrast, emerged as the strongest predictor of NA perception, outperforming melody sensitivity - one of the clearest contributions of the study. This finding seems interesting given that NA is often associated with sentence pitch prominence; thus, the melody test results were expected to play a stronger role. Instead, the finding refines broader claims about the music-language relationship by showing that "musical ability" should not be treated as a unitary construct. Interestingly, rhythm sensitivity did not function as an isolated advantage benefiting all learners equally. The Rhythm effect was remarkably strengthened by higher proficiency scores, particularly in overall and unmarked perception. Auditory-temporal sensitivity seems to be enhanced once learners have sufficient L2 knowledge to interpret what they hear. Rhythm sensitivity may facilitate the detection of prominence, while proficiency helps them assign it linguistic and discourse value. This interaction suggests a layered model of L2 prosodic development, in which auditory sensitivity provides access to cues and proficiency enables learners to integrate those cues into emerging L2 phonological and pragmatic representations.

Unlike in most frequent contexts, no comparable interaction was observed in marked contexts. Marked NA placement appears to impose a higher processing difficulty, as learners not only need to detect prominence but also infer a less transparent relationship between prosody and information structure. Rhythm sensitivity may still stand as a general auditory-attentional advantage in such cases; however, proficiency alone may not suffice to overcome the added demands of leftward prominence shift and de-accentuation. The lack of significant predictors for production is equally revealing, suggesting that NA placement production may hinge on factors beyond the scope of this study such as working memory, inhibitory control, phonological planning, articulatory automatisation, pronunciation learning history, and the quality of previous prosodic instruction.

Considering these findings, English prosody may need to be given more attention in the L2 teaching field, particularly within teacher education programmes, given its central role in understanding meaning and achieving intelligibility, discourse interpretation, and communicative intent. Instruction needs to consider a wider range of intonations patterns than the default pattern of LLI: while this pattern serves as a useful starting point, learners may benefit from more systematic exposure to marked contexts as well, including wh-questions ending in verbs, final adverbials, reflexives, vocatives, and structures involving de-accented tails. The perception and production skills need to be trained separately —learners cannot be expected to produce target-like NA simply because they can identify it, nor should successful controlled production be taken as evidence of perceptual mastery. Effective classroom sequences might include activities to improve NA perception and production such as noticing tasks, metacognitive reflection on cue use, forced-choice identification, visual marking of information focus, and feedback, among other techniques.

Given that rhythm sensitivity predicted NA perception, teachers may find it worthwhile to use rhythm-guided activities as a bridge to prosodic awareness - tapping prominence patterns, comparing stressed and de-accented material, using visual pitch and rhythm displays, marking nuclear accents in transcripts, and contrasting Spanish-like phrase-final prominence with English marked focus patterns. Considering that the L2 language knowledge increases rhythm sensitivity, prosody instruction could reasonably start at early stages of learning rather than be reserved for advanced courses.

One limitation of the current study may be its cross-sectional design which precludes claims about developmental causality. The results show associations among proficiency, musical ability, aptitude, and NA performance, but cannot demonstrate that rhythm sensitivity or proficiency causes improvement in NA perception. Another aspect to be considered is the sample size and their characteristics which in this study was relatively small and specific - Chilean Spanish L1 undergraduate students in English teacher education, all of whom had either taken or were taking an intonation course. This may limit its results generalisation to other Spanish-speaking populations, proficiency levels, instructional contexts, and learner profiles. Further methodological limitations concern the instruments themselves. The proficiency measure offered a practical index of general English level but did not feature extended oral, listening, or discourse-pragmatic measures, and the MLAT subtests used may not have fully captured other aspects of the aptitude construct most relevant to prosody.

Future research might therefore proceed along several lines. Longitudinal and experimental studies are needed to establish whether rhythm-based and information-structure-based instruction can improve NA perception and production over time, and future work could usefully incorporate spontaneous or semi-spontaneous speech tasks alongside controlled utterance production, so that productive control can be examined under more communicatively authentic conditions.

Conclusion

This study examined the perception and production of English NA placement in L2 learners whose first language is Chilean Spanish, together with the role of L2 aptitude, musical ability, and L2 proficiency. The results suggest that these learners' ability to perceive and produce English NA placement was rather limited, yet systematic. Accuracy in both perception and production was shaped by syntactic markedness, prosodic alignment, and task-related conditions, with weaker performance in marked contexts as they require a leftward displacement of nuclear accent instead of the LLI rule. These results underscore the continuing challenge that Spanish L1 learners are faced with when acquiring English NA placement and hig,light the role of cross-linguistic differences in prosodic organisation.

These findings may also contribute to current debates on the perception-production link in L2 phonology. In the current study, production exceeded perception, and neither domain reliably predicted the other. Rather than supporting a simple perception-first account, the results suggest that perception and production may develop separately and recruit distinct processing mechanisms. In this respect, the study adds to evidence that the relationship between the two is non-linear and sensitive to the phonological target, developmental stage, and task demands.

Regarding individual differences, L2 aptitude measured by the MLAT did not predict NA perception or production, suggesting that traditional aptitude measures may be of limited value in explaining prosodic performance in this domain. On the other hand, rhythm sensitivity was found to be a stronger predictor of perceptual accuracy. This effect was shaped by syntactic transparency and strengthened by higher levels of L2 proficiency. The fact that rhythm sensitivity did not predict production may point to a more complex account for productive control, likely involving syntactic integration, phonological planning, and articulatory coordination, in addition to auditory-perceptual sensitivity.

These conclusions should be interpreted cautiously in light of the study's relatively small sample, its focus on broad-focus utterances, and its cross-sectional design. Even so, the study contributes to research on L2 prosody by showing that NA placement is shaped by the interaction of cross-linguistic influence, learner-internal differences, and task demands. More broadly, the findings reinforce the value of giving prosody a more central place in L2 pronunciation research and pedagogy.

ORCID
https://orcid.org/0000-0002-9661-0262
https://orcid.org/0000-0002-7537-1258
https://orcid.org/0000-0001-9977-3510
https://orcid.org/0000-0002-7089-974X
https://orcid.org/0000-0003-0228-3836

Publisher's Note

The claims, arguments, and counter-arguments made in this article are exclusively those of the contributing authors. Hence, they do not necessarily represent the viewpoints of the authors' affiliated institutions, or EUROKD as the publisher, the editors and the reviewers of the article.

Acknowledgements

The authors wish to express their gratitude to the students at the Universidad de Concepción, Chile, for their participation in this study.

Funding

This work was supported by the Chilean National Research and Development Agency, ANID, through Fondecyt grant No. 1220209.

CRediT Authorship Contribution Statement

Christopher Barra-García: Conceptualisation, Writing – Original Draft, Investigation, Data Curation

Mauricio Véliz-Campos: Conceptualisation, Supervision, Writing – Review And Editing, Funding Acquisition

Mauricio A. Figueroa Candia: Supervision, Project Administration

Yasna Ivonne Pereira-Reyes: Supervision, Validation

Mackarena Kartsevski Dune: Formal Analysis

Generative AI Use Disclosure Statement

Regarding the use of generative AI, we confirm that it was not used for content generation. Its use was limited solely to summarising human-authored text and refining the style of the manuscript.

Ethics Declarations

World Medical Association (WMA) Declaration of Helsinki–Ethical Principles for Medical Research Involving Human Participants

The authors explicitly adhered to the World Medical Association Declaration of Helsinki. Participants were provided with written informed consent forms prior to their participation, and their participation was entirely voluntary.

Competing Interests

The authors declare that they have no competing interests.

Data Availability

The datasets generated and analysed during the current study are available from the corresponding author upon reasonable request.

References

Bachman, L., & Palmer, A. (2010). *Language assessment in practice*. Oxford University Press.

Bakkouche, L., & Saito, K. (2026). Effects of auditory processing, memory, and experience on early and later stages of second language speech learning. *Second Language Research, 42*(1), 53-77. https://doi.org/10.1177/02676583251317909

Best, C. (1995). A direct realist view of cross-language speech perception. In W. Strange (Ed.), *Speech perception and linguistic experience: Theoretical and methodological issues* (pp. 171–204). York Press.

Best, C., & Tyler, M. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In M. Munro, & O. Bohn (Eds.), *Language experience in second language speech learning: In honor of James Emil Flege* (pp. 13–34). John Benjamins.

Bidelman, G., Hutka, S., & Moreno, S. (2013). Tone language speakers and musicians share enhanced perceptual and cognitive abilities for musical pitch: Evidence for bidirectionality between the domains of language and music. *PLoS ONE, 8*(4). Article e60676. https://doi.org/10.1371/journal.pone.0060676

Calhoun, S., La Cruz, E., & Olssen, A. (2018). The interplay of information structure, semantics, prosody, and word ordering in Spanish intransitives. *Laboratory Phonology, 9*(1), 1-30. https://doi.org/10.5334/labphon.65

Carroll, J., & Sapon, S. (2002). *Modern language aptitude test (MLAT): Manual (2002 ed.)*. Second Language Testing.

Casillas, J. V. (2019). Phonetic category formation is perceptually driven during the early stages of adult L2 development. *Language and Speech, 63*(3), 550-581. https://doi.org/10.1177/0023830919866225

Casillas, J. V. (2020). The longitudinal development of fine-phonetic detail: Stop production in a domestic immersion program. *Language Learning, 70*(3), 768–806. https://doi.org/10.1111/lang.12392

Cason, N., Marmursztejn, M., D'Imperio, M., & Schön, D. (2020). Rhythmic abilities correlate with L2 prosody imitation abilities in typologically different languages. *Language and Speech, 63*(1), 149–165. https://doi.org/10.1177/0023830919826334

Celce-Murcia, M., Brinton, D., & Goodwin, J. (2010). *Teaching pronunciation: A course book and reference guide* (2nd ed.). Cambridge University Press.

Chalmers, J., Eisenchlas, S. A., Munro, A., & Schalley, A. C. (2021). Sixty years of second language aptitude research: A systematic quantitative literature review. *Language and Linguistics Compass, 15*(11), Article e12440. https://doi.org/10.1111/lnc3.12440

Chobert, J., & Besson, M. (2013). Musical expertise and second language learning. *Brain Sciences, 3*(2), 923–940. https://doi.org/10.3390/brainsci3020923

Dale, P., & Sparks, R. (2024). Mapping the predictive role of MLAT subtests for L2 achievement through regression commonality analysis. *Studies in Second Language Acquisition*, *46*(2), 453–477. https://doi.org/10.1017/S0272263123000530

de Leeuw, E., Stockall, L., Lazaridou-Chatzigoga, D., & Gorba Masip, C. (2021). Illusory vowels in Spanish-English sequential bilinguals: Evidence that accurate L2 perception is neither necessary nor sufficient for accurate L2 production. *Second Language Research*, *37*(4), 587–618. https://doi.org/10.1177/0267658319886623 del Saz, M., & Grau, X. (2022). Production of vowels and nuclear accents by Chilean learners of English as a foreign language. *Revista de Lingüística Teórica y Aplicada*, *60*(1), 219–246. https://doi.org/10.29393/RLA60-8PVMX20008

Delogu, F., & Zheng, Y. (2020). Beneficial effects of musicality on the development of productive phonology skills in second language acquisition. *Frontiers in Neuroscience*, 14, 618–626. https://doi.org/10.3389/fnins.2020.00618

Fadilla, S., Safitri, L., Sakti, G., & Syahrul, S. (2023). The effect of introducing suprasegmental features towards students' listening ability. *Journal of English Language Pedagogy*, *8*(2), 247–261. https://doi.org/10.36665/elp.v8i2.776

Flege, J. E. (1995). Second-language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233–277). York Press.

Flege, J. E. (2003). Assessing constraints on second-language segmental production and perception. In A. Meyer, & N. Schiller (Eds.), *Phonetics and phonology in language comprehension and production: Differences and similarities* (pp. 319–355). Mouton de Gruyter.

Flege, J., & Bohn, O. (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

Frota, S., & Prieto, P. (Eds.). (2015). *Intonation in Romance: Language families, phonetics and phonology*. Oxford University Press.

Ghooch Kanloo, A. H., Kachlicka, M., Saito, K., & Tierney, A. (2026). Individual differences in perception of prosody in Mandarin-accented speech are linked to pitch perception, melody memory, musical training, and neural encoding of sound. *Brain and Language, 273*, Article 105689. https://doi.org/10.1016/j.bandl.2025.105689

Gorba, C., & Cebrian, J. (2021). The role of L2 experience in L1 and L2 perception and production of voiceless stops by English learners of Spanish. *Journal of Phonetics*, *88*(3), Article 101094. https://doi.org/10.1016/j.wocn.2021.101094

Gordon, J., & Darcy, I. (2018). Teaching segmentals vs. suprasegmentals: Different effects of explicit instruction on comprehensibility. *Pronunciation in Second Language Learning and Teaching Proceedings*, *10*(1), 116–126.

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

Götz, A., Liu, L., Nash, B., & Burnham, D. (2023). Does musicality assist foreign language learning? Perception and production of Thai vowels, consonants and lexical tones by musicians and non-musicians. *Brain Sciences*, *13*(5), Article 810. https://doi.org/10.3390/brainsci13050810

Halliday, M. A. K. (1967). *Intonation and grammar in British English*. The Hague.

Hirsch, A., & Wagner, M. (2011, April). *Patterns of prosodic prominence in English intransitive sentences* [Conference presentation]. 34th Generative Linguistics in the Old World (GLOW 34).

Hoot, B., & Leal, T. (2020). Processing subject focus across two Spanish varieties. *Probus*, *32*(1), 93–127. https://doi.org/10.1515/probus-2019-0004

Hosaka, I., & Saito, K. (2026). What promotes success in L2 suprasegmental learning in foreign language classroom settings? The roles of perceptual, cognitive, and experiential individual differences in the acquisition of English word stress. *Applied Psycholinguistics, 47*, Article e5, 1–29. https://doi.org/10.1017/S0142716426100460

Kachlicka, M., Symons, A. E., Ruan, Y., Saito, K., Dick, F., & Tierney, A. (2026). Effects of targeted perceptual training on L2 suprasegmental weighting strategies. *Bilingualism: Language and Cognition*. Advance online publication. https://doi.org/10.1017/S1366728925100850

Kartushina, N., Hervais-Adelman, A., Frauenfelder, U. H., & Golestani, N. (2015). The effect of phonetic production training with visual feedback on the perception and production of foreign speech sounds. *The Journal of the Acoustical Society of America*, *138*(2), 817–832. https://doi.org/10.1121/1.4926561

Kidd, E., Donnelly, S., & Christiansen, M. H. (2018). Individual differences in language acquisition and processing. *Trends in Cognitive Sciences*, *22*(2), 154–169. https://doi.org/10.1016/j.tics.2017.11.006

Kim, S., Y. & Han, J. (2024). The relationship between perception and production of illusory vowels in a second language. *Second Language Research*, *40*(2), 271–300. https://doi.org/10.1177/02676583221135185

Klassen, J. (2013). Second language acquisition of English focus prosody: Evidence from Spanish native speakers. In J. Cabrelli Amaro, T. Judy, & D. Pascual y Cabo (Eds.), *Proceedings of the 12th Generative Approaches to Second Language Acquisition Conference (GASLA 2013* (pp. 76–84). Cascadilla Proceedings Project.

Ladd, D. (1979). Light and Shadow: A study of the syntax and semantics of sentence in English. In L. R. Waugh & F. Van Coetsem (Eds.), *Contributions to grammatical studies: Semantics and syntax*. (pp. 93–131). E. J. Brill.

Ladegaard, H. J., & Chan, K. L. R. (2024). The importance of suprasegmental features in language attitude research: Evidence from a study of teachers' attitudes towards Hong Kong English. *Journal of Multilingual and Multicultural Development*, *46*(10), 3415–3428. https://doi.org/10.1080/01434632.2024.2342923

Landblom, S., A. & Ionin, T. (2022). Nuclear accent placement in broad focus intransitives in native and non-native English: An investigation of syntactic and pragmatic factors. *Glossa: A Journal of General Linguistics*, *7*(1), 1–37. https://doi.org/10.16995/glossa.5810

Lev-Ari, S. (2018). Social network size can influence linguistic malleability and the propagation of linguistic change. *Cognition*, *176*, 31–39. https://doi.org/10.1016/j.cognition.2018.03.003

Levis, J. M., & Wichmann, A. (2015). English intonation: Form and meaning. In M. Reed & J. M. Levis (Eds.), *The Handbook of English pronunciation* (pp. 139–155). Wiley-Blackwell. https://doi.org/10.1002/9781118346952.ch8

Li, C., & Sui, M. (2025). The relationship between complexity, accuracy and fluency in L2 English speech: Individual differences and dynamic patterns. *Studies in Second Language Learning and Teaching*, *15*(3), 501–532. https://doi.org/10.14746/ssllt.28006

Li, M., & DeKeyser, R. (2017). Perception practice, production practice, and musical ability in L2 Mandarin tone-word learning. *Studies in Second Language Acquisition*, *39*(4), 593–620. https://doi.org/10.1017/S0272263116000358

Li, P., Xi, X., & Martin, C. D. (2026). The interaction of auditory acuity and L2 input in suprasegmental cue-weighting shifts after study abroad: Mandarin speakers' perception of Spanish lexical stress. *Second Language Research*. Advance online publication. https://doi.org/10.1177/02676583261441156

Li, S. (2019). Six decades of language aptitude research: A comprehensive and critical review. In Z. Wen, P. Skehan, A. Biedroń, S. Li, & R. L. Sparks (Eds.), *Language aptitude: Advancing theory, testing, research and practice* (pp. 78-96). Routledge. https://doi.org/10.4324/9781315122021-5

Li, S., Hiver, P., & Papi, M. (2022). Individual differences in second language acquisition: Theory, research, and practice. In S. Li, P. Hiver, & M. Papi (Eds.), *The Routledge handbook of second language acquisition and individual differences* (pp. 3–34). Routledge. https://doi.org/10.4324/9781003270546-2

Melnik-Leroy, G. A., Turnbull, R., & Peperkamp, S. (2022). On the relationship between perception and production of L2 sounds: Evidence from Anglophones' processing of the French /u/–/y/ contrast. *Second Language Research, 38*(3), 581–605. https://doi.org/10.1177/0267658320988061

Mennen, I. (2015). Beyond segments: Towards a L2 intonation learning theory. In In E. Delais-Roussarie, M. Avanzi, & S. Herment (Eds), *Prosody and language in contact. Prosody, phonology and phonetics* (pp. 171–188). Springer. https://doi.org/10.1007/978-3-662-45168-7_9

Mennen, I., & de Leeuw, E. (2014). Beyond segments: Prosody in SLA. *Studies in Second Language Acquisition*, *36*(2), 183–194. https://doi.org/10.1017/S0272263114000138

Milovanov, R., Pietilä, P., Tervaniemi, M., & Esquef, P. (2010). Foreign language pronunciation skills and musical aptitude: A study of Finnish adults with higher education. *Learning and Individual Differences*, *20*(1), 56–60. https://doi.org/10.1016/j.lindif.2009.11.003

Nagle, C. L. (2018). Examining the temporal structure of the perception-production link in second language acquisition: A longitudinal study. *Language Learning*, *68*(1), 234–270. https://doi.org/10.1111/lang.12275

Nagle, C. L., & Baese-Berk, M. M. (2022). Advancing the state of the art in L2 speech perception-production research: Revisiting theoretical assumptions and methodological practices. *Studies in Second Language Acquisition*, *44*(2), 580–605. https://doi.org/10.1017/S0272263121000371

Nakabayashi, C. (2021). A case study: A nonnative English teacher teaching suprasegmentals to an EFL learner. *Osaka JALT*, *8*, 61–74.

Nilsenová, M., & Swerts, M. (2012). Prosodic adaptation in language learning. In Romero-Trillo, J. (Eds.), *Pragmatics and prosody in English language teaching*. (pp. 77–96). Springer. https://doi.org/10.1007/978-94-007-3883-6_6

Ortega, L., & Wu, S. L. (2025). Rethinking language proficiency: Commentary on Hulstijn (2024). *Languages*, *10*(4), Article 58. https://doi.org/10.3390/languages10040058

Ortega-Llebaria, M., & Colantoni, L. (2014). L2 English intonation: Relations between form-meaning associations, access to meaning, and L1 transfer. *Studies in Second Language Acquisition*, *36*(2), 331–353. https://doi.org/10.1017/S0272263114000011

Patel, A. (2014). Can nonlinguistic musical training change the way the brain processes speech? The expanded OPERA hypothesis. *Hearing Research*, 308, 98–108. https://doi.org/10.1016/j.heares.2013.08.011

Riquelme, C., Pereira-Reyes, Y., & Véliz-Campos, M. (2025). Unpacking the possible effects of working memory and English proficiency on nuclear accent perception and production in Chilean learners of English. *Language Teaching Research Quarterly*, *49,* 41–58. https://doi.org/10.32038/ltrq.2025.49.03

Romero-Trillo, J. (2012). Teaching prosody with a pragmatic orientation: A synthesis. In J. Romero-Trillo (Ed), *Pragmatics and prosody in English language teaching* (pp. 231–234). Springer. https://doi.org/10.1007/978-94-007-3883-6_14

Şahin Kizil, A. (2024). Linguistic dimensions of L2 performance: Complexity, accuracy and fluency. *Gaziantep Üniversitesi Sosyal Bilimler Dergisi, 23*(4), 1736–1749. https://doi.org/10.21547/jss.1531851

Saito, K. (2021). What characterizes comprehensible and native-like pronunciation among English-as-a-second language speakers? Meta-analyses of phonological, rater, and instructional factors. *TESOL Quarterly*, *55*(3), 866–900. https://doi.org/10.1002/tesq.3027

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., & 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., Petrova, K., Zheng, C., & Tierney, A. (2026). Facets of pitch sensitivity in L2 prosody acquisition: Roles of domain-general and speech-specific individual differences. *Cognition, 273*, Article 106506. https://doi.org/10.1016/j.cognition.2026.106506

Shultz, A. A., Francis, A. L., & Llanos, F. (2012). Differential cue weighting in perception and production of consonant voicing. *Journal of the Acoustical Society of America*, *132*(2), 95–101. https://doi.org/10.1121/1.4736711

Skehan, P. (1998). *A cognitive approach to language learning*. Oxford University Press.

Skehan, P., & Chan, E. Y. S. (2024). Developing an aptitude test based on Pienemann's SLA theory. *Individual Differences in Language Education: An International Journal*, *2*, 1-20. https://doi.org/10.32038/idle.2024.02.01

Sonsaat-Hegelheimer, S., & Levis, J. (2025). Is intonation learnable in the classroom? Evidence from Turkish learners of English. *Language Teaching Research*. Advance online publication. https://doi.org/10.1177/13621688241312502

Sparks, R. (2022). Where does language aptitude come from? Redux. *Language Teaching Research Quarterly*, *31*, 19–40. https://doi.org/10.32038/ltrq.2022.31.03

Sparks, R., & Dale, P. (2023a). The prediction from MLAT to L2 achievement is largely due to MLAT assessment of underlying L1 abilities. *Studies in Second Language Acquisition*, *45*(5), 1345–1369. https://doi.org/10.1017/S0272263123000037

Sparks, R. L., & Dale, P. S. (2023b). Why have L2 educators and researchers largely ignored the importance of L1 skills for L2 aptitude and L2 achievement? *Individual Differences in Language Education: An International Journal*, *1*, 13-28. https://doi.org/10.32038/idle.2023.01.01

Sparks, R., Dale, P., & Patton, J. (2023). Individual differences in L1 attainment and language aptitude predict L2 achievement in instructed language learners. *The Modern Language Journal*, *107*(2), 479–508. https://doi.org/10.1111/modl.12841

Suzukida, Y., & Saito, K. (2022). What is second language pronunciation proficiency? An empirical study. *System*, *106*(1), Article 102754. https://doi.org/10.1016/j.system.2022.102754

Swaminathan, S., Kragness, H. E., & Schellenberg, E. G. (2021). The Musical Ear Test: Norms and correlates from a large sample of Canadian undergraduates. *Behavior Research Methods*, *53*(5), 2007–2024. https://doi.org/10.3758/s13428-020-01528-8

Thompson, R. M., Salig, L. K., & Slevc, L. R. (2025). Is musical ability related to second-language acquisition? A meta-analysis. *Royal Society Open Science*, *12*(1). Article 241193. https://doi.org/10.1098/rsos.241193

Trofimovich, P., Kennedy, S., & Foote, J. A. (2015). Variables affecting L2 pronunciation development. In M. Reed & J. Levis (Eds.), *The handbook of English pronunciation* (pp. 353–373). Wiley Blackwell. https://doi.org/10.1002/9781118346952.ch20

Turker, S., & Reiterer, S. M. (2021). Brain, musicality, and language aptitude: A complex interplay. *Annual Review of Applied Linguistics*, *41*, 95–107. https://doi.org/10.1017/S0267190520000148

Van Maastricht, L., Krahmer, E., & Swerts, M. (2016). Prominence patterns in a second language:

Intonational transfer from Dutch to Spanish and viceversa. *Language Learning*, *66*(1), 124–158. https://doi.org/10.1111/lang.12141
Vercellotti, M. L. (2017). The development of complexity, accuracy, and fluency in second language performance: A longitudinal study. *Applied Linguistics*, *38*(1), 90–111. https://doi.org/10.1093/applin/amv002
Wallentin, M., Højlund, A., Friis-Olivarius, M., & Vuust, P. (2010). The Musical Ear Test, a new reliable test for measuring musical competence. *Learning and Individual Differences*, *20*(3), 188–196. https://doi.org/10.1016/j.lindif.2010.02.004
Wen, Z., Biedroń, A., & Skehan, P. (2017). Foreign language aptitude theory: Yesterday, today and tomorrow. *Language Teaching*, *50*(1), 1–31. https://doi.org/10.1017/S0261444816000276
Yenkimaleki, M. (2019). Explicit teaching of segmentals versus suprasegmentals in developing speaking skills by interpreter trainees: An experimental study. *FORUM*, *17*(2), 249–268. https://doi.org/10.1075/forum.18007.yen
Yenkimaleki, M. (2021). Prosody training benefits in perception vs. production skills in simultaneous interpreting: An experimental study. *Dutch Journal of Applied Linguistics*, *10.* https://doi.org/10.51751/dujal9888
Zubizarreta, M., & Nava, E. (2011). Encoding discourse-based meaning: Prosody vs. syntax: Implications for second language acquisition. *Lingua*, *121*(4), 652–669. https://doi.org/10.1016/j.lingua.2010.06.013