# The perception of Spanish lexical stress by proficient Mandarin learners of Spanish

*Peng Li*[1,2], *Xiaotong Xi*[3]

[1] Basque Center on Cognition, Brain, and Language (BCBL), Spain
[2] University of Oslo, Norway
[3] Shandong University of Finance and Economics, China
pli@bcbl.eu, xiaotong.xi@sdufe.edu.cn

## Abstract

Unlike Spanish natives, Mandarin speakers tend to produce the Spanish lexical stress contrast by manipulating pitch rather than other prosodic cues such as duration. However, the perception of Spanish lexical stress remains less clear. This study examines Mandarin speakers' cue-weighting strategies in perceiving Spanish stress and investigates how musical perception aptitude and auditory processing abilities affect cue-weighting. Twenty-two L1 Mandarin speakers with advanced Spanish proficiency and 19 Spanish natives participated in a stress perception task, which involved identifying strong-weak and weak-strong lexical stress patterns. The pitch and duration ratio of the target word's vowels were manipulated (7 steps each). Musical perception aptitude (i.e., accent, melody, rhythm, and pitch) and auditory processing abilities (i.e., duration and pitch) of Mandarin speakers were also assessed. Results show that Mandarin speakers rely more on pitch than duration to identify Spanish lexical stress patterns, in contrast to Spanish natives (duration > pitch). Additionally, only musical accent perception skills significantly predicted Mandarin speakers' cue-weighting, with higher musical accent scores correlating to larger weighting on pitch cues. The results suggest that even advanced learners exhibit L1 transfer in prosody to L2, and musical perception skills may play a role in L2 prosodic acquisition.

**Index Terms**: lexical stress, musical aptitude, cue-weighting, Spanish, Mandarin, crosslinguistic transfer

## 1. Introduction

Although Spanish lexical stress is largely predictable and sometimes orthographically marked, its acquisition poses challenge to second language (L2) learners, which may lie in correctly identifying the stressed syllables [1].

The importance of word-level suprasegmental cues in the learners' native language (L1) largely determines the cue-weighting in the processing of L2 words [2]. As a consequence, tone L1ers may perform differently from stress L1ers in perceiving L2 lexical tones [3], whereas non-stress L1ers may show less sensitivity to L2 lexical stress than stress L1ers [4]. However, previous studies on the perception of L2 Spanish lexical stress mainly focused on stress L1ers such as English speakers [5], [6], [7], or languages with no lexical stress like Korean [8]. Little research has been done on how tone L1ers perceive Spanish lexical stress, especially on the suprasegmental cues for lexical stress perception. A recent production study showed that proficient L1 Mandarin learners of L2 Spanish manipulated pitch to a larger extent than duration to produce Spanish lexical stress contrasts, which is contrary to Spanish natives who manipulated duration more than pitch [9]. However, no perception study has assessed whether Chinese students rely on pitch more than duration to identify Spanish lexical stress.

Despite learners' L1 background, individual differences also play a role in L2 speech perception [10], among which, musical perception aptitude and domain-general auditory processing abilities [11] have received much attention. First, most of the research focused on the role of musical aptitude in stress L1ers (e.g., English, Spanish) perceiving L2 lexical tones (e.g., Mandarin) [12], [13], [14], [15]. In general, individual's perception aptitude in musical pitch or melody is positively related to their perception performance of L2 lexical tones. Second, the auditory processing abilities of pitch positively correlate to the perception of L2 lexical tones [16]; the temporal (duration and amplitude rising time) and spectral (formant and pitch) processing abilities positively correlate to the perception of L2 contrastive focus [17]. However, little research has focused on how individual differences in the sensitivity of non-speech sound stimuli (e.g., music notes and synthesized sounds) predict the prosodic cues that tone L1ers use to detect L2 lexical stress.

Based on the literature review, this study aims at assessing the following research questions.

RQ1: Do proficient Mandarin learners of Spanish perceive Spanish lexical stress using different cue weighting than Spanish natives? We hypothesize that pitch would be the primary cue for Mandarin speakers to perceive Spanish lexical stress patterns while duration would be secondary (cf. [9] for production data).

RQ2: Do individual differences in musical perception aptitude and auditory processing abilities predict L2 learners' cue-weighting in perceiving Spanish lexical stress? Following previous research [18], [19], we opted four musical components relevant to L2 speech acquisition: accent, melody, pitch, and rhythm. As we are interested in pitch and duration as prosodic cues, we chose duration and pitch from the auditory processing ability test. As previous studies did not reveal many consistent results on L2 lexical stress perception, especially in terms of general auditory processing abilities, RQ2 is largely exploratory, and no specific hypothesis is made.

## 2. Methods

### 2.1. Participants

Twenty-two Mandarin-speaking learners of Spanish (hence force "Chinese students", female = 14, $M_{age}$ = 28.19, ranged

23-35 years old) participated in this study. As proficient late adult learners, the Chinese students were substantially exposed to the target L2. They started their Spanish learning in early adulthood ($M$ =18.32 years old, $SD$ = 1.04). They had received formal Spanish instruction for an average of 4.01 years ($SD$ = 0.5) in China, resided in Spanish-speaking countries for 4.89 years ($SD$ = 2.03), and studies various subjects instructed in Spanish for 2.9 years ($SD$ = 1.60). All Chinese students had successfully passed the DELE test (Diplomas of Spanish as a Foreign Language) at B2 ($n$ = 5), C1 ($n$ = 13), or C2 ($n$ = 4) level, which suggests that their Spanish proficiency ranged from advanced to near native.

As the baseline condition, 19 native speakers of peninsular Spanish ("Spanish natives" thereafter, female = 19, $M_{age}$ = 21, ranged 18-32 years old) participated in the lexical stress perception task. Both groups of participants signed written consent to allow the researchers to process their data.

### 2.2. Materials

#### *2.2.1. Materials for stress perception test*

The stress perception test consisted of a Spanish sentence *Por la plaza [PASO]* 'I pass/he passed through the square', spoken by a female peninsular Spanish speaker (model speaker). The target word was the sentence-final verb *pasar* 'to pass' in two conjugation forms: the first-person singular of the present indicative "*paso* (I pass)" and the third person singular of the preterit indicative "*pasó* (he passed)". The two forms only differ in lexical stress, with *paso* being strong-weak and *pasó* being weak-strong.

We manipulated the duration and mean pitch of the vowels in *pa* and *so* as follows. First, based on previous research indicating a mean pitch difference of 12Hz between stressed and unstressed vowels in female Spanish native speakers [9], we manipulated the mean pitch differences in Hertz between /a/ and /o/ in seven steps (-18, -12, -6, 0, 6, 12, 18), with a 12Hz difference representing the medium-level contrast. Second, considering that female Spanish native speakers produce stressed vowels 1.5 times longer than unstressed vowels [9], we instructed the model speaker to produce *paso* and *pasó* in isolation and found that her mean vowel duration was around 100ms. Hence, we manipulated the vowel duration differences in milliseconds between /a/ and /o/ in seven steps (-75, -50, -25, 0, 25, 50, 75), which resulted in a medium-level duration ratio of 1.5 (150ms:100ms). Finally, the intensity of both syllables was fixed at 64dB, which was the mean intensity of the carrier phrase *Por la plaza*. The synthesis yielded 49 words, and we carefully attached each of the 49 synthesized words after the carrier phrase to create the experimental stimuli.

#### *2.2.2. PROMS-S test battery*

We chose four subsets from the short version of the Profile of Music Perception Skills (PROMS-S) [20]. The 'accent' subtest had musical notes varying in intensity; the 'melody' subtest featured sequences of eighth monophonic notes; the 'pitch' subtest included pure tones varying in pitch; and the 'rhythm' subtest presented two-bar notes with constant intensity but different durations. All subtests aimed to assess participants' perception abilities, ranging from obvious to subtle changes.

#### *2.2.3. Auditory processing test batteries*

We opted two tests from the auditory processing test batteries [21]: duration and pitch. Each subtest comprised a continuum of 100 synthesized stimuli, organized by difficulty, with the test initiating at stimulus level 50 and adjusting difficulty based on participants' responses. The two subsets tested the participants' sensitivity to subtle changes in duration and pitch respectively.

### 2.3. Procedure

All the participants were tested individually in a quiet room with a laptop for displaying the stimuli. No feedback was provided during the entire experiment.

#### *2.3.1. Stress perception test*

The stress perception test was conducted using Praat [22]. First, we presented the participants with written instructions in their native language. Then, the participants proceeded with five exercise trials to familiarize themselves with the experimental procedure. Next, they started the experiment. At each trial, the participants listened to an audio stimulus, which was one of the 49 synthesized stimuli, coupled with an incomplete written sentence on the top of the screen "*Por la plaza, paso/pasó* ___". Two nominative pronouns "*yo* (I)" and "*él* (he)" appeared in the middle of the screen. Participants were instructed to select one of the pronouns to complete the sentence by pressing "Z" or "M" on the keyboard. Hence, if they thought they had heard "*paso* (I pass)", they would choose "*yo* (I)"; otherwise, they would choose "*él* (he)" for "*pasó* (he passed)". The 49 stimuli were repeated four times, resulting in 196 trials. The test lasted around 10 min.

#### *2.3.2. Musical perception test*

The PROMS-S tests were discrimination tasks. Each subtest included 8-10 trials. At each trial, participants listened to the referent stimulus twice, followed by a comparison stimulus. Participants had to indicate whether the comparison stimulus was different from the referent stimulus by selecting one of the five options: *definitely different, probably different, I don't know, probably the same,* and *definitely the same*. The system automatically recorded participants' scores. The test lasted around 15 min.

#### *2.3.3. Auditory processing test*

The auditory processing was an AXB discrimination task. Both duration and pitch subsets had the same structure. At each trial, participants listened to three auditory stimuli and had to indicate whether the second stimulus differed from the first or the third by clicking the numbers "1" or "3" on the screen. The task difficulty was adjusted based on participants' responses. It would be harder after three correct responses and easier after one incorrect response. Participants' score was automatically generated by the system. The test lasted 10 min.

### 2.4. Data coding

First, the participants' responses in the stress perception test were binary coded, with 1 for a *yo* response signifying a strong-weak (SW) pattern and 0 for an *él* response indicating a weak-strong (WS) pattern. The pitch step and duration step were $z$-score transformed for the statistical analysis.

Second, PROMS-S yielded four scores: music accent, music melody, music pitch, and music rhythm, with higher score signifying better musical aptitude. The auditory processing test batteries yielded two scores, which were then transformed to the threshold of perceiving subtle changes in duration (ms) and pitch (Hz): auditory duration, and auditory pitch. Here, a low score means good auditory processing acuity.

To check the predictive role of musical aptitude and auditory processing abilities on cue-weighting, we calculated the participants' cue-weighting on pitch for perceiving Spanish lexical stress, following Yazawa et al. [23]. For each participant, we ran a logistic regression with the response data as the dependent variable, and duration step and pitch step as the independent variables. The coefficients (β) of duration and pitch indicated the extent to which the participants used each cue to perceive Spanish lexical stress. We then calculated the cue-weighting of pitch (pitch cue) with (1). Note that the calculation of the cue-weighting of duration is simply reversed to that of pitch. Hence, pitch cue was set to the dependent variable for statistical analyses.

$$pitch\ cue = \frac{\beta_{pitch}}{\beta_{pitch} + \beta_{duration}} \quad (1)$$

### 2.5. Statistical analyses

The statistical analyses were performed with R software [24]. To answer RQ1, we ran a Generalized Linear Mixed-Effects Model (Model 1) using *lme4* package [25]. The dependent variable was the binary-coded response (1 = SW; 0 = WS). The independent variable included pitch step, duration step, group (Chinese students vs. Spanish natives), and the interaction of Pitch step × Group and Duration step × Group. The reference level was set to Chinese students. Random effects included participant as the random intercept with pitch step and duration step being random slopes.

To answer RQ2, we built two linear regression models (Models 2-3) with pitch cue as the dependent variable. The independent variables were the four PROMS-S scores for Model 2 and the two auditory scores for Model 3. All numeric variables were *z*-score transformed for statistical analyses.

## 3. Results

### 3.1. RQ1: Do proficient Mandarin learners of Spanish perceive the Spanish lexical stress using different cue weighting than Spanish natives?

Fig. 1 shows the results of the lexical stress perception test.

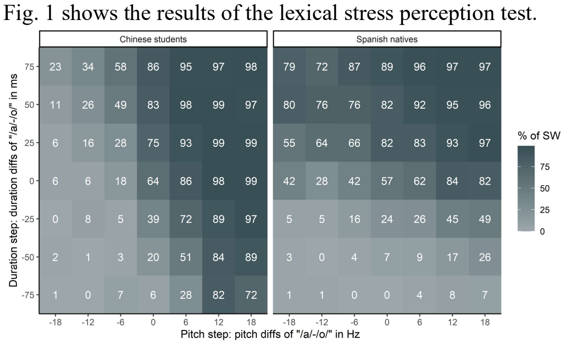


Fig. 1: *Proportion of strong-weak lexical pattern responses to the synthesized stimuli 'PASO' by Chinese students and Spanish natives in the Spanish lexical stress perception test. /a/ is the vowel of the first syllable 'pa'; /o/ is the vowel of the second syllable 'so'.*

Model 1 (Table 1) revealed significant main effects of pitch step and duration step which indicates that an increase in pitch or duration differences between /a/ and /o/ resulted in significantly higher probability for Chinese students choosing SW over WS. The significant Pitch step × Group interaction suggests that the difference between the log-odds ratio corresponding to a change in pitch step for Spanish natives and the log-odds ratio corresponding to that for Chinese students was significant (Fig. 2 P). A similar interpretation applies to the significant Duration step × Group interaction (Fig. 2 D). More importantly, the negative coefficient (-2.23) of Pitch step × Group indicates that when identifying Spanish lexical stress, Spanish natives relied less on pitch than Chinese students did. On the contrary, the positive coefficient (1.11) of Duration step × Group suggests that durational cue was more important for Spanish natives than for Chinese students in identifying Spanish lexical stress patterns. The entire model explained 79% of the variance with the fixed effects accounting for 67.8% of the total variance, which is a large effect size [26].

Table 1: *Summary of Model 1*

| | Fixed effects | | | Random |
|---|---|---|---|---|
| | Log-odds | *SE* | *p* | By participant *SD* |
| (Intercept) | 0.09 | 0.15 | .539 | 0.63 |
| pitch step | 3.21 | 0.22 | <.001 | 0.89 |
| duration step | 1.56 | 0.17 | <.001 | 0.75 |
| group [sp] | -0.07 | 0.21 | .746 | |
| pit step × gr [sp] | -2.23 | 0.30 | <.001 | |
| dur step × gr [sp] | 1.11 | 0.26 | <.001 | |

Model formula:
Response ~ (pitch step + duration step) * group + (pitch step + duration step | participant), reference = Chinese students

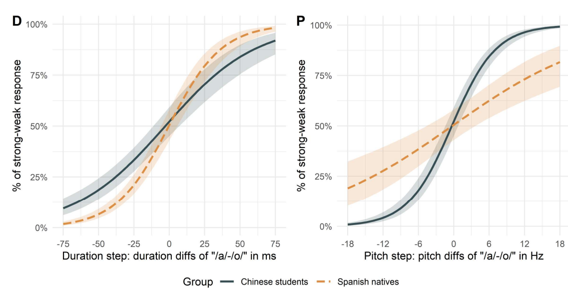


Fig. 2: *Estimated probability of SW response as a function of duration step (Panel D) and pitch step (Panel P). The shaded area indicates 95% CI.*

### 3.2. RQ2: Do individual differences in musical perception aptitude and auditory processing abilities predict L2 learners' cue-weighting in perceiving Spanish lexical stress?

Table 2: *Descriptive data of musical aptitude, auditory processing, and pitch cue of Chinese students*

| | *M* | *SD* | *Min* | *Max* |
|---|---|---|---|---|
| Music aptitude | | | | |

| | | | | |
|---|---|---|---|---|
| Music melody | 5.45 | 2.13 | 2.5 | 9.5 |
| Music rhythm | 3.61 | 1.27 | 0.5 | 5.5 |
| Music accent | 6.09 | 1.84 | 2.5 | 9.0 |
| Music pitch | 3.86 | 1.46 | 2.0 | 7.0 |
| Auditory processing | | | | |
| Auditory pitch (Hz) | 2.67 | 3.01 | 0.95 | 15.7 |
| Auditory duration (ms) | 49 | 25.24 | 16.67 | 92.5 |
| Cue weighting | | | | |
| Pitch cue | 0.66 | 0.11 | 0.45 | 0.88 |

### *3.2.1. Musical perception aptitude*

Model 2 (Table 3) revealed a significant main effect of music accent score, which means that the higher the musical accent score, the more likely the Chinese students rely on pitch to perceive Spanish lexical stress (Fig. 3). The rest of the musical scores did not show any significant effect on cue-weighting.

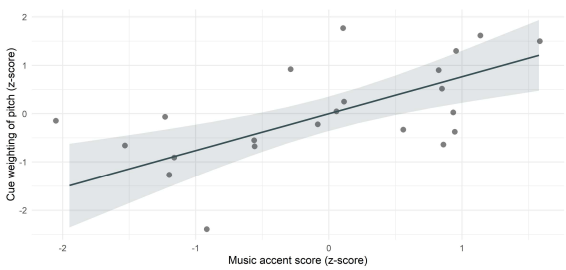


Fig. 3: *Correlation between music accent score and cue-weighting of pitch in L2 Spanish lexical stress perception by Chinese students. The solid line indicates linear functions with shaded area painting 95% confidence interval and scatter plots showing individual data points.*

### *3.2.2. Auditory processing abilities*

Model 3 (Table 3) with auditory processing scores as independent variables did not reveal any significant main effect on Chinese students' pitch cue weighting. This means that auditory processing abilities in duration and pitch are not significant predictors of the extent to which Chinese students use pitch cues to perceive Spanish lexical stress patterns.

Table 3: *Summary of Models 2-3*

| | β | *SE* | *t* | *p* |
|---|---|---|---|---|
| Model 2 | | | | |
| Music melody | -0.14 | 0.21 | -0.65 | .522 |
| Music rhythm | -0.04 | 0.20 | -0.20 | .842 |
| Music accent | 0.76 | 0.19 | 3.98 | .001 |
| Music pitch | -0.35 | 0.20 | -1.73 | .103 |
| Model 3 | | | | |
| Auditory pitch | -0.19 | 0.22 | -0.86 | .401 |
| Auditory duration | -0.17 | 0.22 | -0.74 | .466 |

# 4. Discussion

This study investigated the prosodic cue that proficient L1 Mandarin leaners of L2 Spanish use to identify Spanish lexical stress by a stress perception task. We found that Chinese students rely on pitch more than duration to identify Spanish lexical stress. We further assessed whether musical perception aptitude and auditory processing abilities can predict the L2 learners' cue-weighting of pitch in perceiving Spanish lexical stress and found that only musical accent perception skills positively correlated to the cue-weighting of pitch.

Our data support the hypothesis that pitch is the primary prosodic cue for Chinese students to identify L2 lexical stress and duration is the secondary cue. This finding well aligns with previous production study which showed that the proficient Mandarin learners of Spanish manipulated pitch more than duration to contrast lexical stress [9]. The contrast between Chinese students and Spanish natives provides strong evidence for a crosslinguistic transfer from the learner's L1 to L2 on prosodic level. Because on lexical level, Mandarin varies pitch to contrast word meaning (lexical tone), Mandarin speakers are sensitive to lexical-level pitch specifications, which is carried over to L2 lexical-level prosody. On the other hand, duration is not the key element for Mandarin lexical tone contrast. Therefore, duration is less important than pitch in L2 lexical stress contrast for Chinese students than for Spanish natives.

On individual level, our data confirm that those who are more sensitive to intensity changes in musical stimuli, which was tested by the 'accent' subset of PROMS-S, rely more on pitch to identify L2 lexical stress. This phenomenon can be explained by the pitch-intensity dependence, which holds that the change in intensity may affect listener's perception of pitch height [27]. Therefore, those who are sensitive to intensity change may be sensitive to pitch changes as well. Even though the intensity of the speech stimuli was set to a constant level (64dB), good intensity perceivers may also be good at perceiving pitch. Interestingly, recent cross-linguistic studies showed that tone L1ers with better musical accent perception abilities are better at orally imitating unfamiliar languages whereas non-tone L1ers do not show such correlation [18], [19]. It might well be that musical accent abilities or the sensitivity to intensity play a key role in tone L1ers' L2 speech acquisition, which deserves further investigation.

Nevertheless, the remaining individual factors did not show significant results. Possible reasons are as follows. First, the mean music accent score (6.09) was outstanding among the four musical components, which was not the case of Mandarin speakers in previous studies (e.g., 4.2 in [18]). Therefore, the sampling may have been biased due to unknown reasons although we conducted random sampling. Second, as an initial step of a large project, this study had a relatively small sample size ($N = 22$). This may have hindered the significance of some predictors. Larger-scale study is necessary to advance this topic.

# 5. Conclusions

In conclusion, although with certain limitations, this study demonstrated that the L1-transfer on prosodic level may still occur for advanced L2 learners. More importantly, we provided more evidence to support the close relationship between speech perception and music perception aptitude, which may be used to boost L2 speech learning in teaching practice.

# 6. Acknowledgements

This study is supported by the Research Council of Norway through its Centers of Excellence funding scheme [223265] and the Spanish Ministry of Science, Innovation, and Universities through the Juan de la Cierva program (awarded to PL) funded by "NextGenerationEU"/ PRTR [JDC2022-048729-I].

## 7. References


[1] S. Beaudrie, “La adquisición del acento ortográfico en la clase de español como segunda lengua,” *Hispania*, vol. 90, no. 4, pp. 809–823, 2007, doi: 10.2307/20063614.

[2] A. Cutler, *Native Listening: Language Experience and the Recognition of Spoken Words*. MIT Press, 2012.

[3] A. L. Francis, V. Ciocca, L. Ma, and K. Fenn, “Perceptual learning of Cantonese lexical tones by tone and non-tone language speakers,” *Journal of Phonetics*, vol. 36, no. 2, pp. 268–294, Apr. 2008, doi: 10.1016/j.wocn.2007.06.005.

[4] Z. Qin, Y.-F. Chien, and A. Tremblay, “Processing of word-level stress by Mandarin-speaking second language learners of English,” *Applied Psycholinguistics*, vol. 38, no. 3, pp. 541–570, May 2017, doi: 10.1017/S0142716416000321.

[5] R. Ortín and M. Simonet, “Perceptual Sensitivity to Stress in Native English Speakers Learning Spanish as a Second Language,” *Laboratory Phonology*, vol. 14, no. 1, Jan. 2023, doi: 10.16995/labphon.7978.

[6] R. Ortín and M. Simonet, “Phonological processing of stress by native English speakers learning Spanish as a second language,” *Stud Second Lang Acquis*, vol. 44, no. 2, pp. 460–482, May 2022, doi: 10.1017/S0272263121000309.

[7] J.-Y. Kim, “Perception and Production of Spanish Lexical Stress by Spanish Heritage Speakers and English L2 Learners of Spanish,” in *Selected Proceedings of the 6th Conference on Laboratory Approaches to Romance Phonology*, E. W. Willis, P. M. Butragueño, and E. H. Zendejas, Eds., Somerville, MA: Cascadilla Proceedings Project, 2015, pp. 106–128. [Online]. Available: http://www.lingref.com/cpp/larp/6/abstract3196.html

[8] J. I. Hualde and J. Y. Kim, “The acquisition of Spanish lexical stress by Korean learners,” in *Proceedings of the 18th International Congress of Phonetic Sciences (ICPhS 2015)*, The Scottish Consortium for ICPhS 2015, Ed., London, UK: International Phonetic Association: London, 2015.

[9] P. Li and X. Xi, “Spanish lexical stress produced by proficient Mandarin learners of Spanish,” in *Proceedings of the 4th International Symposium on Applied Phonetics*, Lund, Sweden: ISCA, 2022, pp. 41–45. doi: 10.21437/ISAPh.2022-8.

[10] J. Chobert and M. Besson, “Musical expertise and second language learning,” *Brain Sciences*, vol. 3, no. 2, pp. 923–940, 2013, doi: 10.3390/brainsci3020923.

[11] K. Saito, “How does having a good ear promote successful second language speech acquisition in adulthood? Introducing Auditory Precision Hypothesis-L2,” *Lang. Teach.*, pp. 1–17, 2023, doi: 10.1017/S0261444822000453.

[12] T. L. Gottfried, A. M. Staby, and C. J. Ziemer, “Musical experience and Mandarin tone discrimination and imitation,” *The Journal of the Acoustical Society of America*, vol. 115, no. 5, p. 2545, 2004.

[13] E. Martínez-Montes *et al.*, “Musical expertise and foreign speech perception,” *Frontiers in Systems Neuroscience*, vol. 7, no. NOV, pp. 1–11, 2013, doi: 10.3389/fnsys.2013.00084.

[14] A. Cooper and Y. Wang, “The influence of linguistic and musical experience on Cantonese word learning,” *The Journal of the Acoustical Society of America*, vol. 131, no. 6, pp. 4756–4769, 2012, doi: 10.1121/1.4714355.

[15] M. Li and R. Dekeyser, “Perception practice, production practice, and musical ability in L2 Mandarin tone-word learning,” *Studies in Second Language Acquisition*, vol. 39, no. 4, pp. 593–620, 2017, doi: 10.1017/S0272263116000358.

[16] Z. Qin, R. Jin, and C. Zhang, “The Effects of Training Variability and Pitch Aptitude on the Overnight Consolidation of Lexical Tones,” *J Speech Lang Hear Res*, vol. 65, no. 9, pp. 3377–3391, Sep. 2022, doi: 10.1044/2022_JSLHR-22-00058.

[17] K. Saito, H. Sun, M. Kachlicka, J. R. C. Alayo, T. Nakata, and A. Tierney, “Domain-general auditory processing explains multiple dimensions of L2 acquisition in adulthood,” *Studies in Second Language Acquisition*, vol. 44, no. 1, pp. 57–86, Mar. 2022, doi: 10.1017/S0272263120000467.

[18] P. Li, Y. Zhang, F. Baills, and P. Prieto, “Musical perception skills predict speech imitation skills: Differences between speakers of tone and intonation languages,” *Language and Cognition*, 2023, doi: 10.1017/langcog.2023.52.

[19] P. Li, Y. Zhang, X. Fu, F. Baills, and P. Prieto, “Melodic perception skills predict Catalan speakers’ speech imitation abilities of unfamiliar languages,” in *Proceedings of the 11th International Conference on Speech Prosody*, S. Frota, M. Cruz, and M. Vigário, Eds., 2022, pp. 876–880. doi: 10.21437/SpeechProsody.2022-178.

[20] M. Zentner and H. Strauss, “Assessing musical ability quickly and objectively: development and validation of the Short-PROMS and the Mini-PROMS,” *Annals of the New York Academy of Sciences*, vol. 1400, no. 1, pp. 33–45, 2017, doi: 10.1111/nyas.13410.

[21] I. Mora-Plaza, K. Saito, Y. Suzukida, J.-M. Dewaele, and A. Tierney, “Tools for second language speech research and teaching.” 2022. doi: http://doi.org/10.17616/R31NJNAX.

[22] P. Boersma and D. Weenink, “Praat: doing phonetics by computer [computer program].” 2017. [Online]. Available: http://www.praat.org

[23] K. Yazawa, J. Whang, M. Kondo, and P. Escudero, “Language-dependent cue weighting: An investigation of perception modes in L2 learning,” *Second Language Research*, vol. 36, no. 4, pp. 557–581, Oct. 2020, doi: 10.1177/0267658319832645.

[24] R Core Team, “R: A language and environment for statistical computing.” R Foundation for Statistical Computing, Vienna, Austria, 2014. [Online]. Available: http://www.r-project.org/

[25] D. Bates, M. Mächler, B. Bolker, and S. Walker, “Fitting linear Mixed-Effects Models using {lme4},” *Journal of Statistical Software*, vol. 67, no. 1, pp. 1–48, 2015, doi: 10.18637/jss.v067.i01.

[26] L. Plonsky and H. Ghanbar, “Multiple Regression in L2 Research: A Methodological Synthesis and Guide to Interpreting R2 Values,” *The Modern Language Journal*, vol. 102, no. 4, pp. 713–731, 2018, doi: 10.1111/modl.12509.

[27] C. Arnoldner, A. Kaider, and J. Hamzavi, “The Role of Intensity Upon Pitch Perception in Cochlear Implant Recipients,” *The Laryngoscope*, vol. 116, no. 10, pp. 1760–1765, Oct. 2006, doi: 10.1097/01.mlg.0000228214.02606.42.