



Measuring music and prosody: accounting for variation in non-native speech discrimination with working memory, specialized music skills, and music background

Adam A. Bramlett¹, Bianca Brown², Jocelyn Dueck³, Seth Wiener⁴

Carnegie Mellon University

¹abramlet@andrew.cmu.edu, ²biancab@andrew.cmu.edu, ³jdueck@andrew.cmu.edu, ⁴sethw1@cmu.edu

Abstract

The dynamics of non-native speech perception remain poorly understood, especially in accounting for specialized skills/training. One such skill, musical ability, has been shown to positively impact sensitivity to speech sounds, yet how musical ability is operationalized and measured varies from study to study. Individuals' musical abilities vary in exposure-duration, skill type (e.g., voice, percussion), and skill-level. In the current study, we take an individual differences (n=44) approach to explore sensitivity in non-native speech discrimination of prosodic contrasts. We measure prosodic sensitivity, working memory, and three measures of musical ability: auditory-motor temporal integration [1], auditory discrimination [2, MET]), and musical sophistication [3, Goldsmiths-MSI]. We measured prosodic sensitivity using three AX discrimination tasks and signal detection measures (d'): Mandarin tone (primarily cued by pitch), Italian and Japanese (non-)geminate (primarily cued by duration). Results suggest music background, discrimination, and auditory-motor temporal integration capture related –yet divergent– aspects of music experience. Additionally, music sub-skills (e.g., pitch perception) have unequal contributions to non-native speech sensitivity across languages' respective linguistic cues (e.g., tone). Findings support models of non-native speech perception, which consider cognitive factors and auditory experience outside of language experience.

Index Terms: Individual differences, Music, Non-native speech perception, Measuring prosody

1. Introduction

Non-native speech perception is complex, with many speech contrasts causing difficulties for beginner L2 learners (e.g., [4, 5]). For example, Italian and Japanese learners often struggle to differentiate between geminate and non-geminate contrasts (e.g., Japanese non-geminate /kate/ *win*, geminate /kat:e/ *buying*; Italian non-geminate /ekol/ *echo*, geminate /ɛk:ol/ *here*) [6]. Similarly, Mandarin learners struggle to distinguish between the four tones of Mandarin (e.g., /yu/ means *fish* with a rising tone, *jade* with a falling tone) [7]. Yet, it has long been understood that some individuals come to the task of L2 learning with an advantage, evidenced by consistent variation in proficiencies across language sub-skills, while holding study duration and input constant [8].

One of the most studied areas of individual differences in non-native speech perception concerns L1 transfer effects. The Perceptual Assimilation Model and Speech Learning Model both successfully demonstrate how a learner's L1 influences non-native speech perception and learning [4, 9]. A second but related line of research suggests that individual cues used to discriminate non-native speech contrasts have a major impact

on a learner's ability to acquire speech sounds, particularly for prosodic cues found in geminate and tonal contrasts. For example, [10] suggests that tone experience in the L1 improves L2 tone discrimination. Similarly, Japanese speakers start off with 95% accuracy in Italian geminate discrimination compared to native English speakers with 80% [6]. Further, [11] shows that both duration and sibilant sensitivity transfer in a gradient manner as a function of how much the respective cue is in the L1, regardless of where the cue occurs in the language.

More recently, this same line of individual differences research has looked beyond cue similarities in L1 and L2 by examining the role of music training on L2 speech perception. Linguistic and musical cues often share the same acoustic correlates, e.g., F0 and duration. For example, [12] demonstrates how short term focused musical pitch training is as beneficial for Mandarin tone discrimination as classroom learning. Beyond this, general musical skills/aptitude and even general cognitive abilities (e.g., [8]) have both been linked to and disputed as successful predictors of improved non-native speech perception. For example, [8] suggests that music aptitude is only a weak predictor of variation in non-native speech perception, whereas general cognitive abilities are more reliable. This is in contrast to [13], which finds that musical ability is strongly tied to speech-level language abilities.

These seemingly contradictory results suggest a need for a more nuanced approach to understanding the relationship between individual differences across music subskills and non-native speech discrimination abilities. In the current study, we investigate how different aspects of musical experience predict non-native speech perception by operationalizing music experience in three ways: productive measures (auditory-motor temporal integration) [1], perceptual measures (auditory discrimination) [2, MET]), and musical sophistication [3, Goldsmiths-MSI]. Beyond music, we use a digit span task as a measure of working memory and measure linguistic sensitivity through three AX discrimination tasks: Italian and Japanese geminate contrasts and Mandarin tone contrasts. Italian, Japanese, and Mandarin were chosen because of their straightforward links from acoustic dimensions: geminate contrast → duration; tone contrast → F0. That is, Mandarin sensitivity should be predicted by pitch related skill tasks and Japanese and Italian sensitivity should be predicted by rhythmic skill tasks.

The following three research questions were formulated:

- 1) To what extent do productive and perceptive measures of musical skill predict non-native speech sensitivity in Italian, Japanese, and Mandarin?
- 2) To what extent does self-reported music sophistication predict non-native speech sensitivity in Italian, Japanese, and Mandarin?
- 3) To what extent does working memory predict non-native speech sensitivity in Italian, Japanese, and Mandarin?

2. Methods

2.1. Participants

The recruitment of 48 L1 English speakers was managed through Prolific [14] (n=15) and in-person (n=33) recruitment. An additional 27 potential participants were rejected due to failing initial requirements (i.e., five removed for not being L1 English speakers, eight removed for failed headphone-checks [15], and 14 removed for failed web-camera-checks). We define an L1 English speaker as one who both self-identified as English dominant and acquired English as their first language [16]. To ensure data quality and maximize retained participants, three median absolute deviations (MAD) from median score was calculated as the standard for removal for each task [17]. Of the 48 participants who remained, 4 were removed for low accuracy scores. Of these, 3 were removed for being below MAD range in the language tasks and 1 for low performance in the rhythm auditory-motor integration task. After removal, 44 participants' (age: $\mu = 22.77$, $\sigma = 4.90$) data were retained for analysis.

2.2. Procedure

All experiments, R scripts, and data are available via the Open Science Framework: https://osf.io/bsph6/?view_only=dac11cc3362f419491681fd3364e4da0. Participants took part in a battery of language, music, and domain general tasks using the online experiment builder Gorilla [18]. Participants first completed a two-part headphone check (basic aural attention task and dichotic-pitch task) [15], followed by an eight trial adaptive staircase digit-span task. Next, participants took part in either the music or language segments (order counterbalanced), followed by the complementary segment (i.e., music \rightarrow language, language \rightarrow music). In the language segment, participants took part in three speeded AX-discrimination tasks for Italian, Japanese, and Mandarin (trials automatically proceeded after 1000 ms). All language stimuli were sampled at 44.1 kHz and recorded in sound attenuated booths (Mandarin) or a studio (Japanese and Italian). The Italian and Japanese stimuli and design were taken from [6], and consisted of stop geminate contrasts: Italian - /p t k b d g dʒ/, Japanese - /t k tʃ/. For the Italian AX task, stimuli consisted of 27 pairs of geminate contrasts (e.g., non-geminate /eko/ *echo*, geminate /ek:o/ *here*), which were made up of approximately half real and non-real words, spoken by three native speakers (balanced across position and real/non-word co-occurrence). For the Japanese AX task, stimuli consisted of 33 pairs of geminate contrasts, with approximately half of the geminate and non-geminate pairs matching in pitch-accent (e.g., non-geminate low-high pitch-accent /heta/ *unskilled*, geminate low-high pitch-accent /het:a/ *decreased*) and mismatching with pitch-accent (e.g., non-geminate high-low pitch-accent /kate/ *win*, geminate low-high pitch-accent /kat:e/ *buying*). The Mandarin stimuli and design were taken from [12], and consisted of eight stimuli of the syllable *yu* with four tones (*yu*1, *yu*2, *yu*3, *yu*4) recorded by male and female native speakers. The F0 contours were manipulated using [19]. Each tonal pairing co-occurred equal amounts. That is, each tone occurred with itself three times to equalize match-mismatch answers across the task. In total, there were 52 Italian trials, 66 Japanese trials, and 96 Mandarin trials. In the Italian task, an error led to two pairs only appearing once, which is why there are 52 trials.

The music segment had two basic tasks: auditory-motor temporal integration [1] and auditory discrimination [2, Musical Ear Task], both of which contained melody and rhythm sec-

tions. In the auditory-motor integration tasks, participants heard a rhythm (auditory-motor rhythm) or melody (auditory-motor melody) three times and had to reproduce either the melodic or rhythmic phrase. For example, in the auditory-motor rhythm task each trial played a rhythm three times, with 13 possible beat positions. The participant then needed to repeat the exact beat using the space-bar. Timing was captured on each space-bar press. Similarly, for the auditory-motor melody task a series of seven notes were played. The participant then needed to repeat these with a series of on-screen buttons that corresponded to relative pitches (the melody always started with the middle pitch). After completion of both auditory-motor temporal integration tasks, the participant then completed a rhythm and melody Musical Ear Task. [2], where two auditory music samples (melody or rhythm) were played and then the participant had to determine if the samples were identical (melodies or drum beats) through a button press on the screen.

After completing the music and language segments, participants then filled out a musical sophistication survey [3, Goldsmiths-MSI]. The Goldsmiths-MSI is a self-reported survey, which aims to capture individual differences through musical sophistication in five areas: Active Musical Engagement (e.g., time and money resources spent on music); Self-reported Perceptual Abilities (e.g., musical listening skills); Musical Training (e.g. formal musical training); Self-reported Singing Abilities (e.g. one's own singing); Emotional Engagement with Music (e.g. ability to talk about emotions in music).

2.3. Data analysis

2.3.1. Data wrangling and validation

For reaction time data, the log transformed overall trial reaction time for both the Musical Ear Tasks (rhythm and melody) was calculated. Reaction time trial data outside three median absolute deviations was removed (proportion of trials removed from musical ear tasks: 119 of 5294 trials). Language tasks reaction time were square-root transformed, but no trial data were removed due to the automatic progression of trials at 1000ms.

For accuracy data, a by-item analysis for both Musical Ear Tasks and for each language task was carried out following the same median absolute deviation approach. Retained items by task: Auditory-motor melody items 9/10, Auditory-motor rhythm items 10/10, Mandarin items 8/8, Italian items 24/27, Japanese items 30/33. Additionally in the auditory-motor rhythm task, the first beat of every trial (rhythmic phrase) was removed. This was done because the first beat initiates rhythm capture. As each possible beat is 200 ms apart, the second beat of each rhythmic phrase was then centered by subtracting 100 ms, as seen in Figure 1. After data removal, internal reliability of the items was calculated for each of the musical tasks and questionnaire as tested in the original studies (i.e., Goldsmiths 5 area scores, *Musical Ear Task* melody and rhythm items, auditory-motor melody, auditory-motor rhythm). Table 1 presents Cronbach's alpha values.

For both Musical Ear tasks and language tasks, d-prime was calculated as a measure of sensitivity [20]. For both auditory motor tasks, accuracy was calculated by averaging the percent of beats (rhythm) or pitches (melody) correct on each trial.

2.3.2. Statistical analysis

All quantitative variables from each of the tasks were centered on their mean to 0 and scaled across tasks, as seen in Figure 2. The Goldsmiths General

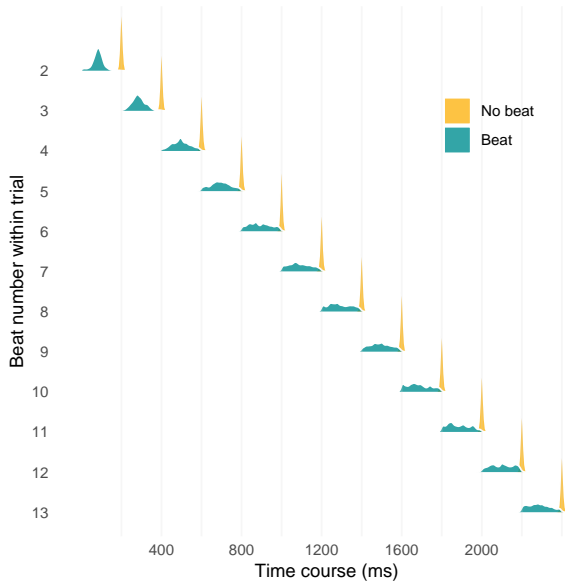


Figure 1: Distribution of beats hit (space bar) over trials. Time frames for each beat (2-13) is illustrated by horizontal lines. No-beat distributions represent [18]’s measurement sensitivity.

Task	Reported α	Our α
GoldSmiths [3]	Range: 0.79 - 0.93	0.89
Musical Ear tasks [2]	0.87	0.79
Auditory-motor melody [1]	Unreported	0.93
Auditory-motor rhythm[1]	Unreported	0.91

Table 1: Task reported Cronbach’s alpha values and our Cronbach’s alpha values

scores were discretized into categories based on three ranges, where score is: Basic: < 36.1 , Average: ≥ 36.1 and < 51.4 , Higher than average: ≥ 51.4 [3].

To answer our three research questions, three multiple linear regression models (one per language) were built to compare the relative contribution of each task on centered and scaled d-prime scores of each language. Maximal models were built using the *lm* package [21] in R [22]. Each model included 10 predictor effects: Goldsmiths engagement, Goldsmiths perception, Goldsmiths training, Goldsmiths singing, Goldsmiths emotions, Musical Ear Test d’ Melody, Musical Ear Test d’ Rhythm, Working-memory, Auditory-motor Rhythm, Auditory-motor Melody. Models were reduced in the following order: removal of non theoretical driven music skill (e.g., rhythm for Mandarin and Melody for Italian and Japanese), then effects were removed based on how much variance was explained in terms of coefficients and post hoc collinearity checks. The Japanese model was reduced minimally due to finding early evidence of parsimonious models. For Italian and Mandarin, the models were reduced to single variables for each theoretically driven question (e.g., music skill, music background) to achieve parsimony and reduce un-needed complexity.

3. Results

Results from the Italian linear model found significant effects for: GoldSmiths emotions ($\beta = -0.23$, $SE = 0.11$, $z = -2.10$, $p <$

0.05), indicating that higher musical emotional ratings predict lower sensitivity to Italian geminate contrasts; Working memory ($\beta = 0.28$, $SE = 0.13$, $z = 2.21$, $p < 0.05$), indicating that higher working memory capacity predicts better ability to attend to sensitivity needed for Italian geminate contrasts.

Results from the Japanese linear model found significant effects for: GoldSmiths active engagement ($\beta = -0.36$, $SE = 0.16$, $z = -2.24$, $p < 0.05$), indicating that higher amounts of musical engagement predict lower sensitivity to Japanese geminate contrasts; Working memory ($\beta = 0.30$, $SE = 0.13$, $z = 2.26$, $p < 0.05$), indicating that higher working memory capacity predicts higher sensitivity to Japanese geminate contrasts; Auditory-motor Rhythm ($\beta = 0.34$, $SE = 0.12$, $z = 2.93$, $p < 0.01$), indicating that higher productive rhythmic ability predicts greater sensitivity to Japanese geminate contrasts.

Results from the Mandarin linear model found significant effects for: GoldSmiths emotions ($\beta = -0.29$, $SE = 0.11$, $z = -2.57$, $p < 0.05$), indicating that higher musical emotional ratings predict lower sensitivity to Mandarin tone; Musical Ear Test melody ($\beta = 0.05$, $SE = 0.05$, $z = 0.05$, $p < 0.05$), indicating that higher perceptual sensitivity to musical pitch predicts higher sensitivity to Mandarin tone.

Figure 3 plots the maximal model terms on the y-axis and coefficient estimates on the x-axis. The visualized estimate (point) and confidence interval (error bar) for each language indicates the results from the parsimonious model. In other words, if a term does not have a present coefficient it means that it was not in the parsimonious model of that language.

4. Discussion

The popular conjecture that musical ability is associated with improved language abilities is not a myth [23], however, the relationship is not as straightforward as the literature may suggest [24]. With respect to our first research question, we found that musical pitch perceptual abilities do indeed account for the most variance in terms of Mandarin tone sensitivity as previously reported in the literature [25]. Similarly, the Japanese model’s results suggest that rhythm ability is indeed tied to Japanese geminate perception in line with previous theoretical claims [26]. Interestingly, motor-auditory integration ability rather than perceptual ability is a superior predictor for sensitivity in Japanese geminate contrasts, while perceptual sensitivity to pitch better accounts for variance in Mandarin tone sensitivity, suggesting that different modalities of musical skill play differential roles across language cues. To the best of our knowledge, this is a new finding not reported in the previous music-to-language transfer literature. This is not to say, that the opposite modalities do not predict linguistic sensitivity, but rather that more variance is accounted for when using pitch-perception to predict Mandarin sensitivity and rhythm-production to predict Japanese sensitivity. Surprisingly, this effect does not carry over to the Italian task. It could be that the difference between Mandarin and Japanese model results have to do with how direct the relationship is between the linguistic cue and the musical cue. In the case of Mandarin tones and musical pitch, the same underlying cue of F0 is used to discriminate both whereas the relationship between geminate contrasts and music rhythm is less clear. Although both rhythm and geminate contrasts use duration cues, we measured the ability to attend to the time between beats or the inter-beat interval, while geminate contrasts often rely on the duration of a consonant. In the case of Italian, it could be that no relationship between rhythm and language sensitivity was found because of the changes in vocalic pronunciation

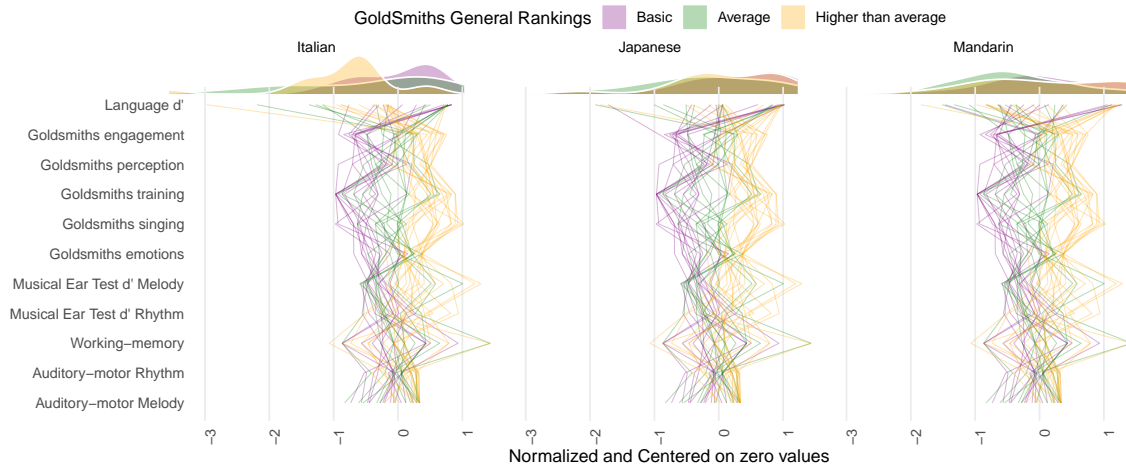


Figure 2: Raw data

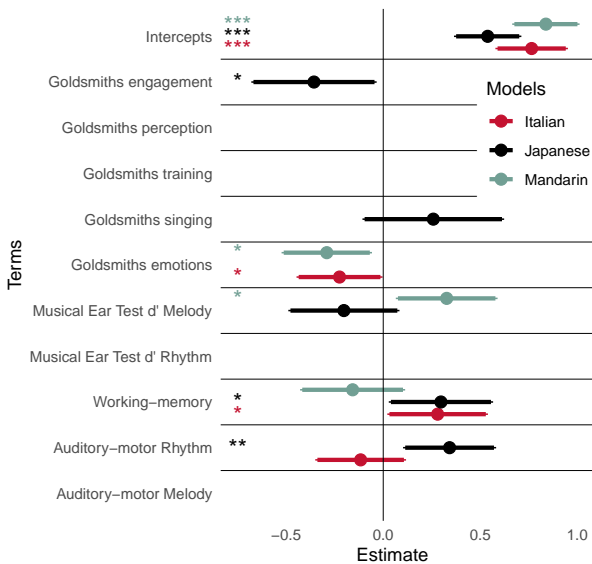


Figure 3: Each language's parsimonious model output

between geminate and non-geminate contrasts (co-articulation) [6]. Thus, Italian geminate sensitivity could be better predicted by the ability to discriminate or produce formant differences.

In terms of research question 2, Table 1 shows that our methods of assessing music ability were reliable. Yet, while specialized skills in music are indeed positively predictive of non-native speech sensitivity, musical self-reported measures have the opposite effect. That is, in all cases throughout the models of each language, higher Goldsmiths music sophistication scores predicted *lower* non-native speech sensitivity. As seen in Figure 2 and corroborated by a post-hoc correlation matrix (see additional OSF materials), Goldsmiths' scores are highly correlated. This unexpected finding could be due to our specific population of university students, the propensity to overestimate one's music ability, or a more general finding that self-reported music ability is not directly associated with any foreign-language ability [27, 28, 29]. We tentatively conclude that higher self-reported music ability predicts lower language

sensitivity, though further research is needed to clarify this finding.

Lastly, with respect to research question 3, both the Japanese and Italian models suggest that working memory is a crucial predictor for sensitivity to geminate contrasts. This finding corroborates recent work that demonstrated how working memory can modulate non-native speech category learning success [30]. Oddly, however, the Mandarin model did not find an effect of working memory. This could be because Mandarin words in our task are single-syllable words based on one syllable, and Japanese and Italian words are varied multi-syllable words. [30] used five different Mandarin syllables and found an effect of working memory; e.g., increased stimuli variation may also increase cognitive demands. This result could also be due to the dimensionality of tone and geminate contrasts. Unlike tone, which has continuous cues for discrimination throughout a word, geminate contrasts happen only in limited segments of Japanese and Italian words. This could mean that partial information in the Mandarin task is sufficient, but higher cognitive demand is necessary for the Japanese and Italian tasks. Regardless, we conclude that domain general cognitive abilities like working memory play a crucial role in speech perception, particularly in the discrimination of non-native prosody, e.g., [1].

5. Conclusions

Our findings indicate that the relationship between music skills, cognitive abilities, and language skills is complex and demands continued investigations. We found that cognitive abilities, perceptual abilities, and productive abilities are variable in predictive power across languages. We also observed unexpected findings in which higher self-rated music abilities predicted weaker sensitivity to non-native speech. Taken together, we call for more nuanced approaches to measuring and assessing music abilities in more varied populations of language users and musicians across different linguistic and musical stimuli.

6. Acknowledgements

This work was supported by a Simon Initiative Seed Grant from Carnegie Mellon University. We thank the instructors who allowed us to recruit in their classes and Kimiko Tsukada and colleagues for sharing their stimuli.

7. References

- [1] M. Kachlicka, K. Saito, and A. Tierney, "Successful second language learning is tied to robust domain-general auditory processing and stable neural representation of sound," *Brain and Language*, vol. 192, p. 15–24, 2019.
- [2] M. Wallentin, A. H. Nielsen, M. Friis-Olivarius, C. Vuust, and P. Vuust, "The musical ear test, a new reliable test for measuring musical competence," *Learning and Individual Differences*, vol. 20, no. 3, p. 188–196, 2010.
- [3] D. Müllensiefen, B. Gingras, J. Musil, and L. Stewart, "The musicality of non-musicians: An index for assessing musical sophistication in the general population," *PLoS ONE*, vol. 9, no. 2, 2014.
- [4] J. E. Flege, "Second language speech learning: Theory, findings, and problems," in *Speech perception and linguistic experience: Issues in cross language research*, W. Strange, Ed. Timonium, MD: York Press, 1995, pp. 229–273.
- [5] J. E. Flege and O. Bohn, "The revised speech learning model (slm-r)," in *Second language speech learning: Theory and empirical progress*, R. Wayland, Ed. Cambridge, UK: Cambridge University Press, 2021, pp. 3–83.
- [6] K. Tsukada, F. Cox, J. Hajek, and Y. Hirata, "Non-native japanese learners' perception of consonant length in japanese and italian," *Second Language Research*, vol. 34, no. 2, p. 179–200, 2017.
- [7] E. Pelzl, E. F. Lau, T. Guo, and R. M. DeKeyser, "Advanced second language learners of mandarin show persistent deficits for lexical tone encoding in picture-to-word form matching," *Frontiers in Communication*, vol. 6, pp. 1–18, 2021.
- [8] C. Zheng, K. Saito, and A. Tierney, "Successful second language pronunciation learning is linked to domain-general auditory processing rather than music aptitude," *Second Language Research*, vol. 38, no. 3, 2021.
- [9] C. T. Best, "A direct realist perspective on cross-language speech perception," in *Speech perception and linguistic experience: Theoretical and methodological issues in cross-language speech research*, W. Strange, Ed. Timonium, MD: York Press, 1995, pp. 167–200.
- [10] 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, p. 268–294, 2008.
- [11] B. Pajak and R. Levy, "The role of abstraction in non-native speech perception," *Journal of Phonetics*, vol. 46, pp. 147–160, 2014.
- [12] S. Wiener and E. D. Bradley, "Harnessing the musician advantage: Short-term musical training affects non-native cue weighting of linguistic pitch," *Language Teaching Research*, vol. 27, no. 4, p. 1016–1031, 2020.
- [13] L. R. Slevc and A. Miyake, "Individual differences in second-language proficiency: Does musical ability matter?" *Psychological Science*, vol. 17, no. 8, 2006.
- [14] S. Palan and C. Schitter, "Prolific.ac—a subject pool for online experiments," *Journal of Behavioral and Experimental Finance*, vol. 17, p. 22–27, 2018.
- [15] A. E. Milne, R. Bianco, K. C. Poole, S. Zhao, A. J. Oxenham, A. J. Billig, and M. Chait, "An online headphone screening test based on dichotic pitch," *Behavior Research Methods*, vol. 53, no. 4, p. 1551–1562, 2021.
- [16] B. Brown, B. Tusmagambet, V. Rahming, C.-Y. Tu, M. B. DeSalvo, and S. Wiener, "Searching for the "native" speaker: A preregistered conceptual replication and extension of Reid, Trofimovich, and O'Brien (2019)," *Applied Psycholinguistics*, vol. 44, no. 4, p. 475–494, 2023.
- [17] C. Leys, C. Ley, O. Klein, P. Bernard, and L. Licata, "Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median," *J. Exp. Soc. Psychol.*, vol. 49, no. 4, pp. 764–766, Jul. 2013.
- [18] A. L. Anwyl-Irvine, J. Massonnié, A. Flitton, N. Kirkham, and J. K. Evershed, "Gorilla in our midst: An online behavioral experiment builder," *Behavior Research Methods*, vol. 52, no. 1, p. 388–407, 2019.
- [19] P. Boersma and D. Weenink, "Praat: doing phonetics by computer [computer program]." [Online]. Available: <https://www.fon.hum.uva.nl/praat/>
- [20] N. Macmillan and C. Creelman, *Detection Theory: A User's Guide*, 2nd ed. Psychology Press, 2004.
- [21] R. C. Team, *lm: Linear Models in R*, R Foundation for Statistical Computing, Vienna, Austria, 2021. [Online]. Available: <https://stat.ethz.ch/R-manual/R-devel/library/stats/html/lm.html>
- [22] R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2021. [Online]. Available: <https://www.R-project.org/>
- [23] L. R. Slevc and A. Miyake, "Individual differences in second-language proficiency," *Psychological Science*, vol. 17, no. 8, p. 675–681, 2006.
- [24] L. Neves, A. I. Correia, S. L. Castro, D. Martins, and C. F. Lima, "Does music training enhance auditory and linguistic processing? a systematic review and meta-analysis of behavioral and brain evidence," *Neuroscience & Biobehavioral Reviews*, p. 104777, 2022.
- [25] Y. Zheng and A. G. Samuel, "The effects of ethnicity, musicianship, and tone language experience on pitch perception," *Quarterly Journal of Experimental Psychology*, vol. 71, no. 12, pp. 2627–2642, 2018.
- [26] A. Löfqvist, "An effect of rhythm class?" in *The Phonetics and Phonology of Geminant Consonants*, 2017, vol. 2, p. 118.
- [27] P. Larrouy-Maestri, X. Wang, R. V. Nunes, and D. Poeppel, "Are you your own best judge? on the self-evaluation of singing," *Journal of Voice*, vol. 37, no. 4, pp. 633–e7, 2023.
- [28] A. I. Correia, C. F. Lima, and E. G. Schellenberg, "Self-awareness of musical ability," *Psychology of Aesthetics, Creativity, and the Arts*, 2023, advance online publication.
- [29] E. G. Schellenberg, A. I. Correia, and C. F. Lima, "Is musical expertise associated with self-reported foreign-language ability?" *Journal of Experimental Psychology: Human Perception and Performance*, 2023.
- [30] J. R. McHaney, R. Tessmer, C. L. Roark, and B. Chandrasekaran, "Working memory relates to individual differences in speech category learning: Insights from computational modeling and pupilometry," *Brain and Language*, vol. 222, p. 105010, 2021.