NON-NATIVE TONE PERCEPTION – WHEN MUSIC OUTWEIGHS LANGUAGE EXPERIENCE

Antonia Götz [1,2] & Liquan Liu [1,3,4,5]

1 The MARCS Institute for Brain, Behaviour and Development, Western Sydney University, Sydney, Australia,
2 Department of Linguistics, University of Potsdam, Potsdam, Germany,
3 School of Psychology, Western Sydney University, Sydney, Australia,
4 Center of Multilingualism across the Lifespan, University of Oslo, Oslo, Norway,
5 Centre of Excellence for the Dynamics of Language, Australian Research Council, Canberra, Australia.

ABSTRACT

This study examined how language (e.g., bilingualism, L2) and music (e.g., years of practising) experiences improve lexical tone perception. A substantial number of 532 participants from L1 Mandarin, L1 non-tone, bilingual L1 non-tone & L2 non-tone, and bilingual L1 non-tone & L2 tone backgrounds were tested on their discrimination of Mandarin tones. Results revealed that neither bilingual nor second (tone or non-tone) language experience affects novel tone perception. However, listeners' years of music training predicted perception outcomes regardless of listeners' language backgrounds. These results indicate 1) learning a tone language as L2 does not guarantee perceptual advantage of non-native tones, even after years of learning; 2) a myth of "bilingual advantage" in tone perception, challenging the bilingual enhanced acoustic sensitivity hypothesis in the perceptual domain; and 3) learning a musical instrument helps with tone perception across language groups, exhibiting a cross-domain effect in the processing of linguistic and musical pitch.

Keywords: tone perception, bilingualism, second language perception, years of musical training, cross-domain effect.

1. INTRODUCTION

In tone languages, tones distinguish word meanings. Around 60-70% of the world's languages are tonal [1], and more than half of the world population speak a tone language [2]. Most tone perception studies have been conducted on individuals from tonal languages, such as Mandarin or Cantonese, and there is a limited understanding of how individuals from non-tonal languages perceive tone and how second language learning assist in perceiving lexical tones. This project extended from existing tone perception studies and investigated whether tone experience learned from one's second language, sequential bilingual experience and music experience play a role in tone perception. Specifically, the study can inform the development of more effective language learning strategies and interventions for individuals with tone perception difficulties, as well as shed light on the cognitive benefits of bilingualism and musical training.

Some tones or tone contrasts are naturally more discriminable than others. In Mandarin tones, T1-T3 (level-dipping) appears more salient than T2-T3 (rising-dipping) for tone and non-tone language speakers [3], and T2-T4 (rising-falling) exceeds other contrasts in discriminability [4]. In Cantonese tones,

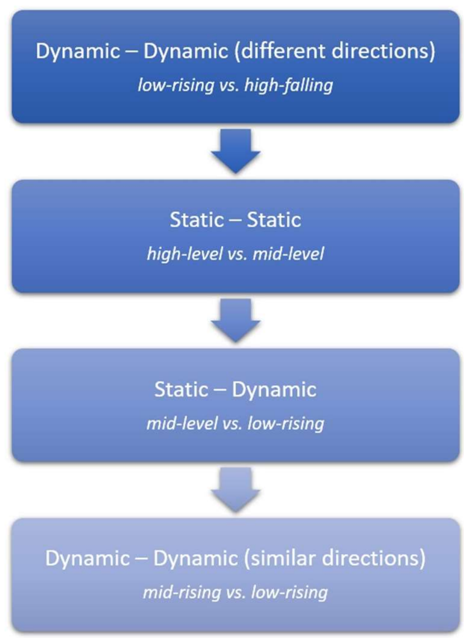

Figure 1: In [6], participants from five language backgrounds (Australian English, Cantonese, Chinese Mandarin, Singaporean Mandarin, Thai) unanimously show the same pattern when perceiving static and dynamic tones across four tone languages (Cantonese, Chinese Mandarin, Singaporean Mandarin, Thai). A salience hierarchy is hypothesised marking which type of tone contrast is easier (higher) to discriminate than others (lower) regardless of listeners' language backgrounds.

T2-T6 (high rising-low rising) leads to lower discrimination accuracy than other contrasts across listeners from tone and non-tone, monolingual and bilingual backgrounds [5]. A recent collaborative study involving participants from five tone and non-tone language backgrounds reports a salience hierarchy regarding the type of tone contrast (Figure 1). Prior research has established that tone language speakers perceive tones more categorically than their non-tone language peers [7]. What remains less clear is whether tone language speakers perceive non-native tones as well as native ones, and to what extent listeners' L2 tonal experience would lead to successful tone discrimination. One recent study testing Cantonese, Chinese, Singaporean and Thai listeners' perception of tones from their own versus other languages reveals that tone language speakers do not perceive non-native tones as good as their native ones [6]. When testing Mandarin perception by L2 Mandarin learners from L1 Cantonese or English backgrounds, the two groups do not differ in their overall accuracy on Mandarin tones, and both groups show difficulty with the acoustically most difficult T2-T3 tone pair in Mandarin [8]. Tone perception appears to be difficult even for advanced L2 learners [9], and listeners' L1 knowledge and the acoustic features of L2 tones both interfere with tone perception.

With respect to bilingualism, simultaneous bilingual experience has been shown to strengthen tone perception in infancy [10], [11] and adulthood [12] even when tone is not part of the language repertoires of the bilingual speakers. Such perceptual advantage is often attributed to the bilingual environment typically more complex than a monolingual one. The need of establishing two phonological systems may well enhance listeners' sensitivity to a third. Whether such advantage extends to the cognitive domain is under heated debate [13]. As speakers' L2 experience can also be considered as a sequential bilingual experience, it is interesting to see whether (non-tone) L2 experience may play a role, as well as whether the length of L2 learning may be a relevant factor in this case.

In terms of music experience, music training has been shown to improve non-tone language speakers' tone perception [14]–[16]. Listeners with no prior tone language experience discriminate tones more accurately when they are more musically trained [17], [18]. Some studies suggest that non-tone language speakers perceive (linguistic) tones in the same fashion of musical tunes, as their performances of pitch in language and music correlate [19]. These findings point to a domain-general effect showing that music experience can enhance non-tone language speakers' tone perception. Bilinguals and tone language speakers do not show such correlation [12], suggesting the interference of bilingual and L1 experience on tone perception. The research questions of the current study are: 1. Do different tone contrasts lead to different perceptual outcomes? 2. Does (L1 or L2) tone language or sequential bilingualism experience strengthen novel tone perception? Do listeners' years of L2 experience play a role? 3 Does music experience (years of music training) modulate tone perception? We predicted that based on the salience hierarchy [6], Mandarin T2-T4 (dynamic different) would be the most salient contrast, followed by T1-T2/T3/T4 (static-dynamic), and then T3-T2/T4 (dynamic similar). Moreover, listeners' L2 (tone, non-tone) and music experiences facilitate tone perception, and the magnitude of facilitation is directly relevant to their years of experience.

2. METHODS

2.1. Participants

A total number of 814 subjects participated in the study. Of these participants, 191 did not complete the online experiment and another 91 participants did not provide valid consent sharing data for analysis. The final sample included 532 participants from multiple language backgrounds who were divided into three language groups (Table 1).

Group	N	M_{age} (range)	Years of L2 (mean (SD))	Years of music (mean (SD)
L1 Mandarin	25	18-25	12.65 (4.55)	2.93 (2.91)
Mono (non-tone L1)	44	18-25	0.48 (0.09)	2.83 (2.93)
Bi (non-tone L1 + non-tone L2)	429	18-25	8.97 (5.14)	3.26 (3.53)
Bi (non-tone L1 + tone L2)	34	18-25	9.85 (4.94)	3.09 (3.05)

Table 1: Participants were categorised into three groups. Here, Mono refers to monolinguals and Bi refers to sequential bilinguals.L1 Mandarin speakers had L2 experience with non-tone languages but not in another tone langauge. Years of music refers to the mean and standard deviation of years of musical training that was self-indicated by the participants.

2.2. Stimuli

The stimuli consisted of 12 monosyllabic Mandarin non-words (/tou/, /bou/, /ɕye/, /pye/, /pian/, /fian/, /jy/, /ty/, /bi/, /gi/, /gua/, /lua/) with legal phonotactic structures. Each syllable was produced with the four Mandarin tones (T1, T2, T3 and T4). The length of each syllable was 250 ms. The final stimulus set consisted of 72 stimuli: 12 syllables x 6 tone contrasts (T1-T2, T1-T3, T1-T4, T2-T3, T2-T4, T3-T4). A Mandarin native speaker produced six tokens of each syllable and we included two of the tokens into the experiment. By including different tokens, we prevent that the participants make their decision based on acoustic information alone. The same token was never repeated within one trial of the task. For example, in AAB trials, if token 1 was used as the A sound, token 2 would appear as the sound for X (the second A. In addition to the variability of the different tokens, the same word was never repeated within one Tone Contrast. All stimuli were normalized in intensity (70 dB).

2.3. Procedure

The experiment ran online by using the online platform Labvanced [20] at a quiet place and participants were asked to wear headphones. Each experiment started with two practice trials to familiarise the participants with the AXB discrimination task. Participants were asked to press a key as accurate and quick as possible if the second syllable was more similar to the first one (AAB, via key 1) or the third (ABB, via key 3) syllable. All contrast were also presented in the reverse order to counterbalance participants' responses. In each trial, the A and B sounds consisted of a different tone category. The interstimulus interval was 1000 ms and the intertrial interval was 3000 ms. The time-out of the response time was set to 2500 ms, measured at the end of the third syllable. Trials with responses after the time-out were not repeated. The next trial started immediately following the previous trial and trials were randomised across participants. A break was included after 25%, 50% and 75% of the experimental trials and its length was participant controlled; participants continued the experiment by pressing a key. No feedback was provided to the participants.

3. RESULTS

All statistical analyses were performed by using R [21] and the lme4 package [22]. Plots were generated by using ggplot2 [23]. General Linear Mixed Effects regression models were constructed with the maximal random and fixed factor structure with accuracy (binomial answer as 1 or 0) as dependent variable. Language background was coded as the comparison between L1 native Mandarin (coded as 0.5) and monolingual L1 (coded as -0.5), monolingual L1 (coded as 0.5) and bilingual L2 Non-tone speakers (coded as -0.5) and bilingual L2 Non-tone speakers (coded as 0.5) versus L2 Tone speakers (coded as -0.5).

Within the Tone Contrast discrimination, we predicted the following hierarchy: T2-T4 > T1-T2/T3/T4 > T3-T2/T4. Following this hierarchy, we applied the subsequent contrast comparisons: T2-T4 was compared to the mean of T1-T2, T1-T3 and T1T4, T3-T4 was compared to the mean of T1-T2, T1-T3 and T1T4, T3-T4, T3-T4 was compared to T2-T3, T1-T3 was compared to T1-T4 and T1-T2 was compared to T1-T3. Listener's discrimination of the Tone Contrast showed the expected patterns: T2-T4 yielded higher accuracy than T1-TN contrasts (β(SE) = 0.461 (0.019), z = 23.289, p < .001), which yielded higher accuracies than T3-T4 (β(SE) = 0.165(0.014), z = 11.957, p < .001), which was further higher than T2-T3 ($\beta(SE)$ = 0.254(0.024), z = 10.390, p < .001).

Our results (see also Figure 2) revealed that L1 Mandarin speakers are not better in discriminating Tone Contrasts than L1 Non-tone speakers ($\beta(SE)$ = -0.031(0.065), z = 0.478, p = 0.633). L2 tone language speakers are not better in discriminating Tone Contrasts than L2 Non-tone language speakers ($\beta(SE)$ = -0.043(0.041), z = -1.054, p = 0.291) and the listener's years of experience did not interact ($\beta(SE)$ = 0.035(0.070), z = 0.498, p = 0.619). Similarly, L2 listeners of a non-tone language were not better in discriminating the Tone Contrasts than L1 Non-tone language listeners ($\beta(SE)$ = -0.020(0.034), z = -0.594, p = 0.553), and duration of the second language did not interact ($\beta(SE)$ = -0.034(0.070), z = 0.483, p = 0.629).

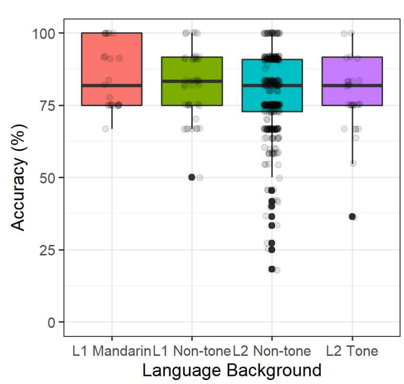

Figure 2: Accuracy results from the speech discrimination task separated by the four groups of L1 and L2 tone language experience (Mandarin, Monolingual L1 Non-tone, Bilingual L2 Non-tone, Bilingual L2 Tone)

However, music experience modulated Tone Contrast discrimination. The more music experience listeners have, the better their tone discrimination ($\beta(SE)$ = 0.011(0.005), z = 2.077, p = 0.0383).

4. DISCUSSION

The current study investigated the extent to which L2 (tone or non-tone) language and music experiences modulate tone perception. With respect to tone contrasts, a salience hierarchy as predicted was reported. Contrasting our predictions, listeners' L1 and L2 language experience does not alter tone perception, regardless of whether the language is tonal or how many years they have learned the language. However, music experience plays an important role for speakers across language backgrounds, with positive relationship between years of music training and successful tone perception outcomes.

In terms of Mandarin tone contrasts, dynamic tones with different pitch directions (T2-T4, rising-falling) were the easiest to discriminate, followed by static-dynamic tone contrasts (T1-T2/T3/T4, flat-contour tones) and then dynamic tones with similar pitch patterns (T3-T2/T4, dipping-rising/falling). Our findings replicate the salience hierarchy reported in Liu and colleagues [6], suggesting that tone acoustics is critical for listeners' tone perception across language backgrounds.

Learning a tone language as L2 does not lead to successful tone perception. Previous studies have reported similar findings for L1 speakers of a non-tone and even a tone language learning another tone language as L2 [8]. Even advanced L2 learners may find it difficult to (re-)establish a tone category [9], conforming to our results that years of L2 experience does not affect tone perception outcomes. The finding is in line with another study on segmental contrast perception. Catalan has an /e/–/ɛ/ contrast while Spanish has only one /e/ closer to and more open than the Catalan /e/. Pallier and colleagues [24] examined two types of Catalan-Spanish bilinguals' perception of the Catalan contrast. The Catalan-dominant bilinguals are exposed to Catalan since birth, whereas the Spanish-dominant bilinguals are exposed to Spanish first and get in touch with Catalan after 6 years of age when they start in kindergarten or primary school. Only Catalan-dominant but not Spanish-dominant bilinguals discriminate the Catalan contrast. Results of the vowel perception among Spanish-dominant bilinguals are in accordance with the findings of tone perception in the current study. The overall findings also demonstrate the importance of early exposure on phonological category establishment and later perception, mirroring studies reporting perceptual advantages [25], [26] and neural traces [27] of phonology of the birth language among adoptees who are adopted to a new language environment as early as 6 months after birth.

Although a simultaneous bilingual experience has been shown to facilitate tone perception [10], a sequential one does not appear to reach a similar effect. While Dutch sequential bilinguals show reduced tone discrimination abilities than their Mandarin peers, Dutch simultaneous bilinguals stand in the middle between these two groups in their performances [12]. The overall findings add to the limited data of how sequential bilingual may impact speech perception, suggesting the importance of early linguistic diversity on tone perception. The lack of significant results between the language groups may be attributed to the differences in group sizes and the resulting larger variability in Bilingual L2 Non-tone group.

Regarding music experience, the current results conform to and extend on prior studies [14]–[16]. That is, listeners do not have to be musicians to show an advantage in tone perception. Their years of music training is directly relevant to tone perception irrespective of their language backgrounds. The finding that overall music experience can facilitate speech perception provides new insights into potential cross-domain strategies in L2 language learning, especially when the target language is tonal.

5. CONCLUSION

When perceiving non-native tones, listeners' music experience plays an important role. Its effect may be more evident than L2 experiences, as neither L2 tone language nor sequential bilingual experience appears to affect perception of foreign tones. The overall findings provide implications for early, consistent diversity in and exposure to language and music.

6. REFERENCES

[1] M. Yip, *Tone*. Cambridge University Press, 2002.

[2] V. A. Fromkin, *Tone: A linguistic survey*. Academic Press, 2014.

[3] Z. Zeng, L. Liu, A. Tuninetti, V. Peter, F.-M. Tsao, and K. Mattock, "English and Mandarin native speakers' cue-weighting of lexical stress: Results from MMN and LDN," *Brain and Language*, vol. 232, p. 105151, 2022.

[4] T. Huang and K. Johnson, "Language specificity in speech perception: Perception of Mandarin tones by native and nonnative listeners," *Phonetica*, vol. 67, no. 4, pp. 243–267, 2010.

[5] X. Tong, S. M. K. Lee, M. M. L. Lee, and D. Burnham, "A tale of two features: Perception of Cantonese lexical tone and English lexical stress in Cantonese-English bilinguals," *PloS one*, vol. 10, no. 11, p. e0142896, 2015.

[6] L. Liu *et al.*, "The tone atlas of perceptual discriminability and perceptual distance: Four tone languages and five language groups," *Brain and Language*, vol. 229, p. 105106, 2022.

[7] P. A. Hallé, Y.-C. Chang, and C. T. Best, "Identification and discrimination of Mandarin Chinese tones by Mandarin Chinese vs. French listeners," *Journal of Phonetics*, vol. 32, no. 3, pp. 395–421, 2004.

[8] Y.-C. Hao, "Second language acquisition of Mandarin Chinese tones by tonal and non-tonal language speakers," *Journal of Phonetics*, vol. 40, no. 2, pp. 269–279, Mar. 2012, doi: 10.1016/j.wocn.2011.11.001.

[9] E. Pelzl, E. F. Lau, T. Guo, and R. DeKeyser, "Advanced second language learners' perception of lexical tone contrasts," *Studies in Second Language Acquisition*, vol. 41, no. 1, pp. 59–86, 2019.

[10] L. A. Petitto, M. S. Berens, I. Kovelman, M. H. Dubins, K. Jasinska, and M. Shalinsky, "The 'Perceptual Wedge' hypothesis as the basis for bilingual babies' phonetic processing advantage: New insights from fNIRS brain imaging," *Brain Lang*, vol. 121, no. 2, pp. 130–143, May 2012, doi: 10.1016/j.bandl.2011.05.003.

[11] L. Liu and R. Kager, "Perception of tones by bilingual infants learning non-tone languages," *Bilingualism: Language and Cognition*, vol. 20, no. 3, pp. 561–575, 2017.

[12] L. Liu, A. Chen, and R. Kager, "Simultaneous bilinguals who do not speak a tone language show enhancement in pitch sensitivity but not in executive function," *Linguistic Approaches to Bilingualism*, vol. 12, no. 3, pp. 310–346, 2022.

[13] M. Lehtonen, A. Soveri, A. Laine, J. Järvenpää, A. De Bruin, and J. Antfolk, "Is bilingualism associated with enhanced executive functioning in adults? A meta-analytic review.," *Psychological bulletin*, vol. 144, no. 4, p. 394, 2018.

[14] D. Burnham, R. Brooker, and A. Reid, "The effects of absolute pitch ability and musical training on lexical tone perception," *Psychology of Music*, vol. 43, no. 6, pp. 881–897, Nov. 2015, doi: 10.1177/0305735614546359.

[15] 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.

[16] P. C. Wong and T. K. Perrachione, "Learning pitch patterns in lexical identification by native English-speaking adults," *Applied Psycholinguistics*, vol. 28, no. 04, pp. 565–585, 2007.

[17] J. A. Alexander, P. C. M. Wong, and A. Bradlow, "Lexical tone perception in musicians and non-musicians," presented at the European Conference on Speech Communication and Technology, Lissabon, Portugal, 2005.

[18] B. Chandrasekaran, A. Krishnan, and J. T. Gandour, "Relative influence of musical and linguistic experience on early cortical processing of pitch contours," *Brain Lang*, vol. 108, no. 1, pp. 1–9, Jan. 2009, doi: 10.1016/j.bandl.2008.02.001.

[19] A. Chen, L. Liu, and R. Kager, "Cross-domain correlation in pitch perception, the influence of native language," *Language, Cognition and Neuroscience*, vol. 31, no. 6, pp. 751–760, 2016.

[20] H. Finger, C. Goeke, D. Diekamp, K. Standvoß, and P. König, "LabVanced: a unified JavaScript framework for online studies," in *International conference on computational social science (cologne)*, 2017.

[21] 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/

[22] D. M. Bates, Maechler, M., Bolker, B., and Walker, S., "Fitting Linear Mixed-Effects Models using lme4," *Journal of Statistical Software*, 2015.

[23] H. Wickham, "Programming with ggplot2," in *ggplot2: Elegant Graphics for Data Analysis*, H. Wickham, Ed. Cham: Springer International Publishing, 2016, pp. 241–253. doi: 10.1007/978-3-319-24277-4_12.

[24] C. Pallier, L. Bosch, and N. Sebastián-Gallés, "A limit on behavioral plasticity in speech perception," *Cognition*, vol. 64, no. 3, pp. B9–B17, 1997.

[25] W. Choi, X. Tong, F. Gu, X. Tong, and L. Wong, "On the early neural perceptual integrality of tones and vowels," *Journal of Neurolinguistics*, vol. 41, pp. 11–23, 2017.

[26] W. Choi, X. Tong, and S. H. Deacon, "Double dissociations in reading comprehension difficulties among Chinese–English bilinguals and their association with tone awareness," *Journal of Research in Reading*, vol. 40, no. 2, pp. 184–198, 2017.

[27] L. J. Pierce, J.-K. Chen, A. Delcenserie, F. Genesee, and D. Klein, "Past experience shapes ongoing neural patterns for language," *Nature Communications*, vol. 6, no. 1, pp. 1–11, 2015.