Interspeech 2018
2-6 September 2018, Hyderabad

Learning two tone languages enhances the brainstem encoding of lexical tones

Akshay Raj Maggu[1], Wenqing Zong[1], Vina Law[2], Patrick C. M. Wong[1]

[1]The Chinese University of Hong Kong
[2]University of Toronto

akshay@cuhk.edu.hk, zongwenqing@cuhk.edu.hk, vina.law@mail.utoronto.ca, p.wong@cuhk.edu.hk

Abstract

Auditory brainstem encoding is influenced by experience-dependent factors such as language and music. Tone language speakers exhibit more robust brainstem encoding of lexical tones than non-tone language speakers. Studies suggest that the effects of experience with a tone language generalize to the brainstem encoding of lexical tones from other tone languages. However, the effects of learning two tone languages, with different tonal systems, on brainstem encoding of lexical pitch are unknown. In the current study, we investigated whether or not the experience with two tone languages (Mandarin and Cantonese) enhances the brainstem encoding of lexical pitch, using frequency following response (FFR). Mandarin has four lexical tones- high level, rising, dipping, and falling while Cantonese has a richer tone system with three level tones (high, mid, low), two rising tones (high and low), and one falling tone. We compared speakers fluent in Cantonese vs. those fluent in both Cantonese and Mandarin on their brainstem encoding of Cantonese and Mandarin lexical tones. We found that the Cantonese-Mandarin speakers exhibited more robust brainstem encoding of the lexical tones as compared to Cantonese speakers. From the current findings, we conclude that learning two tone languages may enhance lexical pitch encoding at the brainstem.

Index Terms: tone language, bilingualism, brainstem, frequency following response

1. Introduction

Previous studies suggest that experience-dependent effects of language and music are facilitatory for language perception [1]–[5]. Recent behavioral evidence suggests that a combination of the experience-dependent effects of language and music are not additive and thus, does not result in enhanced linguistic perception [6]. However, whether or not a combination of experience-dependent effects of different languages are additive is an intriguing question. In order to understand this, we compared subjects with experience of two tone languages with those with experience of only one tone language, in relation to brainstem encoding of lexical tones using a Frequency Following Response (FFR).

Auditory brainstem processing is influenced by experience-dependent factors such as language and/or music [1], [7]–[10]. Sound encoding at the level of the brainstem can be studied using FFR that enables examination of the phase-locked neural response to the incoming sound stimulus. Fidelity of the FFR to the stimulus determines the magnitude of encoding at the level of the brainstem. In order to study the effects of language experience on brainstem encoding, tone languages make a suitable case as they contain lexical tones that are pitch contours carrying different meanings. For example, the syllable /ma/ in Mandarin Chinese means 'mother' when spoken with a high level tone (T1), 'hemp' when spoken with a rising pitch contour (T2), 'horse' when spoken with a dipping tone (T3), and 'to scold' when spoken with a falling pitch contour (T4). Lexical tones are useful in investigating brainstem encoding as the FFR is phase-locked to the stimulus fundamental frequency, the acoustic correlate of tone [11]. Previous studies, using FFR, reveal that tone language speakers exhibit a more robust brainstem encoding of lexical tones than non-tone language speakers [7]–[9]. Further, Krishnan et al. [8] found that tone language speakers can generalize to the brainstem encoding of lexical tones from other tone languages. For example, their study revealed that Mandarin speakers had similar brainstem encoding of Thai lexical tones to that of native Thai speakers.

FFR studies from bilingual listeners of non-tone languages reveal that bilingualism enhances the brainstem encoding of speech sounds, presumably via corticofugal modulation [12]–[15]. However, until now, there have been no studies that have investigated the effect of the combination of experiences with two tone languages on the brainstem encoding of lexical pitch contours. Behavioral [6] and neural findings [16] from previous studies suggests that the effect of the combination of two experiences (language and music) is no better than one. In other words, auditory experiences that are otherwise facilitatory for language perception, might end up causing a "saturation effect" when in combination.

In the current study, we aimed at understanding whether or not the effects of learning two tone languages are additive at the level of the brainstem. Using FFR measures, we compared speakers who are fluent in both Cantonese and Mandarin to those who are fluent in Cantonese (but not Mandarin), on the brainstem encoding of Cantonese and Mandarin lexical tones. The lexical tones used in the current study were *rising* (T2; occurs in both Cantonese and Mandarin), *dipping* (T3m; occurs in Mandarin but not in Cantonese), and *low-falling* (T4c; occurs in Cantonese but not in Mandarin). If the effects of learning two tone languages turned out to be additive, we would predict that the Cantonese-Mandarin group would exhibit enhanced brainstem encoding in comparison to the Cantonese group, similar to the findings from the studies conducted in bilingualism [12]–[15]. On the other hand, if experience with two tone languages do not result in added advantages, as noted in the literature from behavioral studies [6], we would expect a similar magnitude of brainstem encoding of lexical tones across the two groups.

10.21437/Interspeech.2018-2130

2. Method

2.1. Participants

All participants (N = 43; 14 males) were native Hong Kong Cantonese speakers recruited from the Chinese University of Hong Kong. They had peripheral hearing sensitivity within 25dB HL for the frequencies 0.5 to 4 kHz, with no history of middle ear pathology and/or obvious anatomical defects of the speech and hearing mechanism. The participants with a self-reported proficiency of "7" on a 7-point rating scale for both Cantonese and Mandarin were included in the Cantonese-Mandarin (CM) group while those who self-reported "7" for Cantonese and less than "3" for Mandarin were included in the Cantonese group. We obtained 20 participants (mean age = 20.95 years; 6 males) in the Cantonese group and 23 participants (mean age = 21.30 years; 8 males) in the Cantonese-Mandarin group. The groups were matched based on their musical experience, education, socio-economic status, and IQ. The proficiency status of the participant groups was confirmed by 24 native Mandarin-speaking raters on four questions that dealt with the proficiency and usage of Mandarin in daily life (Figure 1). The current study was approved by The Joint Chinese University of Hong Kong – New Territories East Cluster Clinical Research Ethics Committee.

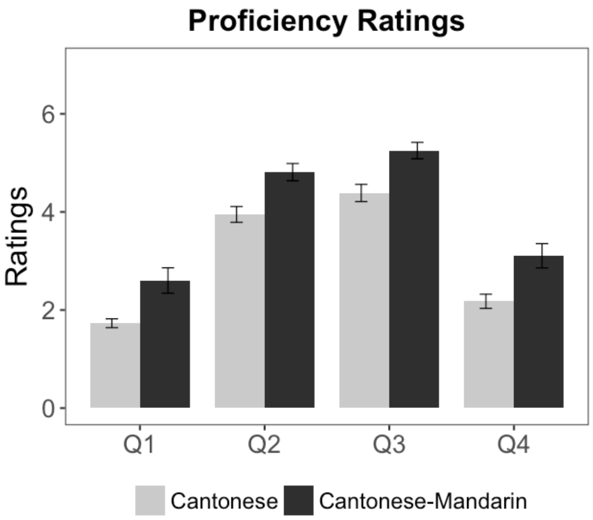

Figure 1: *Comparison of mean ratings of Mandarin proficiency of Cantonese and Cantonese-Mandarin groups by 24 native Mandarin raters (Error bars = ± SEM). Cantonese-Mandarin group were rated to have a better Mandarin proficiency than the Cantonese group.*

2.2. Stimuli

The stimuli consisted of the syllable /ga/ recorded with the three lexical tones (T2, T3m, T4c) from Cantonese and Mandarin making three unique words. The stimuli were intensity-normalized to 75 dB SPL and time-normalized to 175 ms using Praat [17]. Figure 2 shows the F0 contours of the three lexical tones.

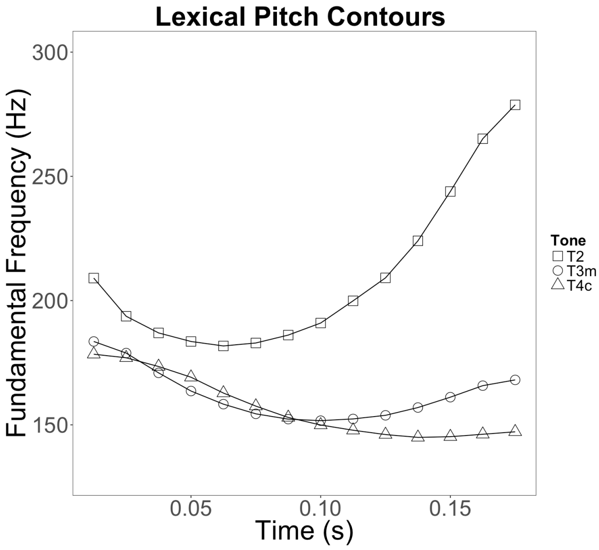

Figure 2: *Lexical pitch contours – T2 (rising; range F0 = 180 – 278 Hz), T3m (dipping; range F0 = 151 – 184 Hz), and T4c (low-falling; range F0 = 147 – 178 Hz).*

2.3. Procedure

2.3.1. Stimuli Presentation

A total of 3000 sweeps of each stimulus were presented binaurally in alternating polarity to each subject via insert earphones (Compumedics 10Ω) at 81 dB SPL using the Audio CPT module of STIM2 (Compumedics Neuroscan, USA). The inter-stimulus (offset to onset) interval was jittered from 74 to 104 ms [1], [11], [16], [18] and the order of stimulus presentation was randomized across participants. Participants were asked to relax and ignore the stimuli.

2.3.2. Data acquisition and pre-processing

Continuous electrophysiological data were collected at a sampling rate of 20 kHz using Ag/AgCl electrodes at Cz (active) referenced to linked M1 and M2 (linked mastoids) with the lower forehead as the ground. The inter-electrode impedances were maintained at ≤ 1 kΩ. A pre-processing pipeline consisting of baseline correction (-50 ms), artifact rejection (± 35 µV), filtering (80-5000 Hz; 6 dB roll off), epoching, and averaging was conducted using Curry Neuroscan 7.05 (Compumedics, El Paso, TX). FFR recordings with not more than 10% rejected sweeps (i.e., > 300 rejections) were included in the current study.

2.3.3. FFR pitch analyses

The pre-processed data were further band-pass filtered to 80-2500 Hz to attenuate high-frequency noise in the signal. As we were interested in the pitch of the FFR signal, these data were converted to the frequency domain from the time domain to extract the F0 contour using a 40-ms sliding window Fast Fourier Transform in 1 ms steps [1], [18], [19]. The data were further processed to compare the Cantonese-Mandarin group and the Cantonese group across the following three measures [16], [18], [20]–[22].

1. Stimulus-to-response correlation (ranges from -1 to +1): It is a simple correlation between the pitch contour of the

stimulus and the FFR signal. Higher positive values of stimulus-to-response correlation reflect better brainstem encoding.

2. Pitch Strength (ranges from 0 to 1): This was obtained by an autocorrelation technique and is a measure of periodicity of the FFR signal. Higher values of pitch strength mean better brainstem encoding.
3. Pitch Error (in Hz): This refers to the average Euclidean distance between the stimulus and FFR pitch (F0) contours. The lower the pitch error, the better the pitch encoding at the brainstem.

3. Results

Figure 3 shows the grand averaged running autocorrelograms (a measure of pitch strength) of the Cantonese-Mandarin and Cantonese groups. Overall, the Cantonese-Mandarin group demonstrated stronger pitch encoding as compared to the Cantonese group for all the three lexical tones. A series of 2 (Group: Cantonese-Mandarin vs. Cantonese) × 3 (Tone: 3 lexical tones) ANOVAs were conducted for each of the FFR measures.

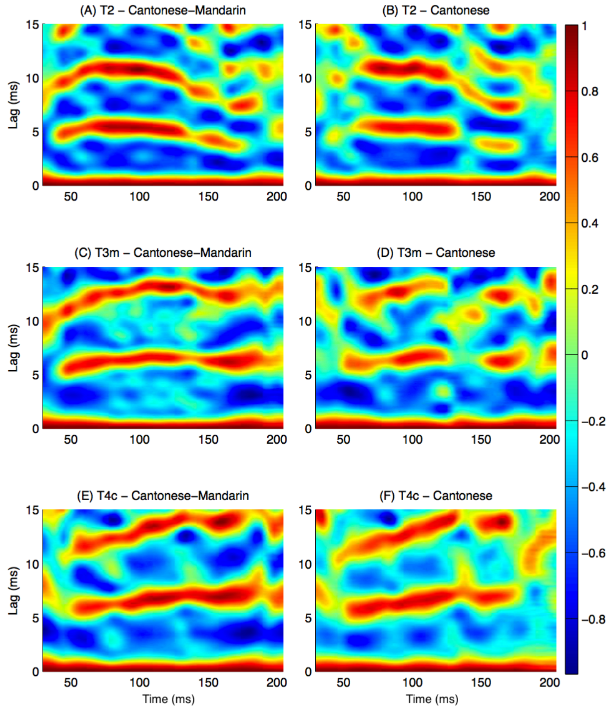

Figure 3: *Comparison of Cantonese-Mandarin (A, C, E) and Cantonese (B, D, F) groups on the running autocorrelogram measure of pitch tracking of lexical tones. The warmer the color, the better the brainstem encoding. The Cantonese-Mandarin group exhibited better pitch-tracking than the Cantonese group on all the three lexical tones, i.e., T2 (A vs. B), T3m (C vs. D), and T4c (E vs. F).*

For stimulus-to-response correlation, there was a main effect of group (F(1, 41) = 18.15, p = 0.000), no main effect of tone ((2, 82) = 1.13, p = 0.339) and no group × tone interaction (F(2, 82) = 1.09, p = 0.339). For pitch strength, there was a main effect of group (F(1, 41) = 5.52, p = 0.024), main effect of tone (F(2, 82) = 9.76, p = 0.000) but no group × tone interaction (F(2, 82) = 2.91, p = 0.06). For pitch error, there was a main effect of group (F(1, 41) = 6.58, p = 0.014), main effect of tone (F(2, 82) = 4.68, p = 0.012) but no group × tone interaction (F(2, 82) = 1.03, p = 0.361). Overall, the Cantonese-Mandarin group outperformed the Cantonese group on all the three FFR measures (Figure 4).

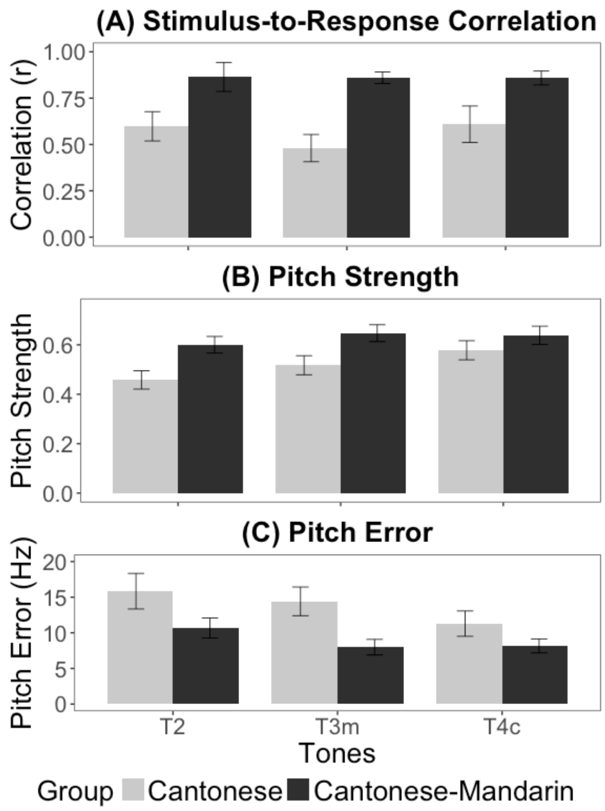

Figure 4: *Comparison of Cantonese and Cantonese-Mandarin groups across the three FFR measures - (A) Stimulus-to-Response Correlation; (B) Pitch Strength; and (C) Pitch Error. The Cantonese-Mandarin group outperformed the Cantonese group on all the three FFR measures (Error bars = ±SEM).*

4. Discussion

In the current study, we investigated the effects of the combination of experience-dependent effects of two tone languages on brainstem encoding of lexical pitch. Using FFR, we found that the participants who were experienced with Cantonese and Mandarin exhibited enhanced brainstem encoding of both Cantonese and Mandarin lexical tones compared to those who had experience of Cantonese alone.

The current findings are in contrast to the behavioral findings of Cooper and Wang [6] who found a "saturation" of experience-dependent effects on lexical tone perception. This could probably be due to the effects of different types of auditory experiences (language and music) involved in their study. Though language and music are both facilitatory for linguistic perception, the mechanisms of language and music perception are speculated to differ in the case of tone language speakers [6]. In comparison, in the current study, a similar type of auditory experience (language) is involved, probably as a result of which we found the effects of experience of two tone languages to be additive. The current findings are in agreement with the previous findings [12]–[15] from non-tone languages that suggest that bilinguals exhibit enhanced brainstem encoding of speech sounds. Bilinguals are known to have a higher gray matter density in the brain regions underlying communication [23], [24]. Krizman et al. [13] speculate that there may be possible higher gray matter density in the inferior colliculi, a generator site for FFR, in bilinguals compared to monolinguals which leads to enhanced brainstem encoding. Further, recent findings suggest that the effects visible at the brainstem could originate at the level of the cortex and be driven to the auditory brainstem via corticofugal pathways [12], [25]–[29]. In the current study, the Cantonese-Mandarin speakers would have had more overall combined experience of different tonal systems than Cantonese speakers which could have led to their enhanced encoding at the level of the brainstem via corticofugal modulation. Another explanation for the current findings could be from the predictive tuning model [30] according to which there is continuous online modulation of the brainstem from the cortex via corticofugal pathways [31], [32] with the local processes at the inferior colliculi still active. Signal representation is enhanced when there is a match between the incoming stimulus and the prediction from the cortical level. It is possible that Cantonese-Mandarin speakers, with overall greater experience of perceiving the lexical tone stimuli, could have a more enhanced signal at the level of the brainstem due to more faithful matching to the incoming stimuli as compared to Cantonese speakers.

5. Conclusion

Similar to bilingualism of two non-tone languages, we found evidence of enhanced brainstem encoding in bilingual tone-language speakers. Our results suggest that long-term experience-dependent neuroplasticity is extended to similar auditory experiences.

6. Future directions

In order to ascertain the generalizability of the current findings, future studies could be conducted with speakers of tone languages other than Cantonese and Mandarin. In addition, future studies could also consider investigating other event-related potentials.

7. Acknowledgements

This research was supported by the Dr Stanley Ho Medical Development Foundation and Research Grants Council of Hong Kong Grant No.14117514 to P.C.M.W.

8. References

[1] P. C. M. Wong, E. Skoe, N. M. Russo, T. Dees, and N. Kraus, “Musical experience shapes human brainstem encoding of linguistic pitch patterns,” *Nat. Neurosci.*, vol. 10, no. 4, pp. 420–422, Apr. 2007.

[2] P. C. M. Wong and T. K. Perrachione, “Learning pitch patterns in lexical identification by native English-speaking adults,” *Appl. Psycholinguist.*, vol. 28, no. 04, pp. 565–585, 2007.

[3] J. A. Alexander, P. C. M. Wong, and A. R. Bradlow, “Lexical tone perception in musicians and non-musicians.,” in *Interspeech*, 2005, pp. 397–400.

[4] G. M. Bidelman, J. T. Gandour, and A. Krishnan, “Cross-domain Effects of Music and Language Experience on the Representation of Pitch in the Human Auditory Brainstem,” *J. Cogn. Neurosci.*, vol. 23, no. 2, pp. 425–434, Nov. 2009.

[5] G. M. Bidelman, J. T. Gandour, and A. Krishnan, “Musicians and tone-language speakers share enhanced brainstem encoding but not perceptual benefits for musical pitch,” *Brain Cogn.*, vol. 77, no. 1, pp. 1–10, Oct. 2011.

[6] A. Cooper and Y. Wang, “The influence of linguistic and musical experience on Cantonese word learning,” *J. Acoust. Soc. Am.*, vol. 131, no. 6, pp. 4756–4769, 2012.

[7] A. Krishnan, Y. Xu, J. Gandour, and P. Cariani, “Encoding of pitch in the human brainstem is sensitive to language experience,” *Cogn. Brain Res.*, vol. 25, no. 1, pp. 161–168, Sep. 2005.

[8] A. Krishnan, J. T. Gandour, and G. M. Bidelman, “The effects of tone language experience on pitch processing in the brainstem,” *J. Neurolinguistics*, vol. 23, no. 1, pp. 81–95, 2010.

[9] A. Krishnan, J. T. Gandour, G. M. Bidelman, and J. Swaminathan, “Experience dependent neural representation of dynamic pitch in the brainstem,” *Neuroreport*, vol. 20, no. 4, pp. 408–413, Mar. 2009.

[10] J. Swaminathan, A. Krishnan, and J. T. Gandour, “Pitch encoding in speech and nonspeech contexts in the human auditory brainstem,” *Neuroreport*, vol. 19, no. 11, pp. 1163–1167, Jul. 2008.

[11] A. R. Maggu, F. Liu, M. Antoniou, and P. C. M. Wong, “Neural Correlates of Indicators of Sound Change in Cantonese: Evidence from Cortical and Subcortical Processes,” *Front. Hum. Neurosci.*, vol. 10, 2016.

[12] J. Krizman, V. Marian, A. Shook, E. Skoe, and N. Kraus, “Subcortical encoding of sound is enhanced in bilinguals and relates to executive function advantages,” *Proc. Natl. Acad. Sci.*, vol. 109, no. 20, pp. 7877–7881, 2012.

[13] J. Krizman, E. Skoe, V. Marian, and N. Kraus, “Bilingualism increases neural response consistency and attentional control: Evidence for sensory and cognitive coupling,” *Brain Lang.*, vol. 128, no. 1, pp. 34–40, Jan. 2014.

[14] J. Krizman, E. Skoe, and N. Kraus, “Bilingual enhancements have no socioeconomic boundaries,” *Dev. Sci.*, vol. 19, no. 6, pp. 881–891, 2016.

[15] E. Skoe, E. Burakiewicz, M. Figueiredo, and M. Hardin, “Basic neural processing of sound in adults is influenced by bilingual experience,” *Neuroscience*, vol. 349, pp. 278–290, May 2017.

[16] A. R. Maggu, P. C. Wong, M. Antoniou, O. Bones, H. Liu, and F. C. Wong, “Effects of combination of linguistic and musical pitch experience on subcortical pitch encoding,” *J. Neurolinguistics*, vol. 47, pp. 145–155, 2018.

[17] P. Boersma and D. Weenink, “Praat, a system for doing phonetics by computer,” *Glot Int.*, vol. 5, no. 9/10, pp. 341–345, 2001.

[18] F. Liu, A. R. Maggu, J. C. Y. Lau, and P. C. M. Wong, “Brainstem encoding of speech and musical stimuli in congenital amusia: evidence from Cantonese speakers,” *Front. Hum. Neurosci.*, vol. 8, p. 1029, 2015.

[19] J. H. Song, E. Skoe, P. C. M. Wong, and N. Kraus, “Plasticity in the adult human auditory brainstem following short-term linguistic training,” *J. Cogn. Neurosci.*, vol. 20, no. 10, pp. 1892–1902, Oct. 2008.

[20] E. Skoe and N. Kraus, “Auditory brain stem response to complex sounds: A tutorial,” *Ear Hear.*, vol. 31, no. 3, pp. 302–324, Jun. 2010.

[21] J. H. Song, K. Banai, and N. Kraus, “Brainstem timing deficits in children with learning impairment may result from corticofugal origins,” *Audiol. Neurotol.*, vol. 13, no. 5, pp. 335–344, May 2008.

[22] P. C. M. Wong, E. Skoe, N. M. Russo, T. Dees, and N. Kraus, “Musical experience shapes human brainstem encoding of linguistic pitch patterns,” *Nat. Neurosci.*, vol. 10, no. 4, pp. 420–422, 2007.

[23] V. Ressel *et al.*, “An effect of bilingualism on the auditory cortex,” *J. Neurosci. Off. J. Soc. Neurosci.*, vol. 32, no. 47, pp. 16597–16601, Nov. 2012.

[24] A. Mechelli *et al.*, “Neurolinguistics: structural plasticity in the bilingual brain,” *Nature*, vol. 431, no. 7010, p. 757, Oct. 2004.

[25] N. Suga, “Role of corticofugal feedback in hearing,” *J. Comp. Physiol. A*, vol. 194, no. 2, pp. 169–183, 2008.

[26] A. Parbery-Clark, E. Skoe, and N. Kraus, “Musical experience limits the degradative effects of background noise on the neural processing of sound,” *J. Neurosci.*, vol. 29, no. 45, pp. 14100–14107, 2009.

[27] V. M. Bajo, F. R. Nodal, D. R. Moore, and A. J. King, “The descending corticocollicular pathway mediates learning-induced auditory plasticity,” *Nat. Neurosci.*, vol. 13, no. 2, pp. 253–260, 2010.

[28] J. H. Song, E. Skoe, K. Banai, and N. Kraus, “Training to improve hearing speech in noise: biological mechanisms,” *Cereb. Cortex*, vol. 22, no. 5, pp. 1180–1190, 2012.

[29] E. Skoe and N. Kraus, “Musical training heightens auditory brainstem function during sensitive periods in development,” *Front. Psychol.*, vol. 4, 2013.

[30] B. Chandrasekaran, E. Skoe, and N. Kraus, “An integrative model of subcortical auditory plasticity,” *Brain Topogr.*, vol. 27, no. 4, pp. 539–552, 2014.

[31] 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, 2009.

[32] B. Chandrasekaran and N. Kraus, “The scalp-recorded brainstem response to speech: Neural origins and plasticity,” *Psychophysiology*, vol. 47, no. 2, pp. 236–246, Mar. 2010.