

The Relationship Between L2 Phonological Processing and Early Musical Training in Adult Bilinguals

Music & Science
Volume 3: 1–11

© The Author(s) 2020

Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/2059204320919285

journals.sagepub.com/home/mnsAngela Liang  and Marcus Taft

Abstract

Bilinguals who speak English as their second-learned (L2), yet dominant language from a young age, have been previously shown to experience difficulties in adulthood when dealing with spoken English (i.e., phonological processing) compared to native monolinguals. The present study investigated whether the processing disadvantage in this group of bilinguals could be modulated by engagement in early musical training, a practice that, like phonological processing, requires the complex manipulation of sounds. To this end, performance on three English auditory processing tasks (phoneme deletion, spelling-to-dictation, and auditory comprehension) was measured in early L2-dominant bilingual and native English monolingual adults who had or had never received early musical training. The processing difficulties that were detected in bilingual adults without musical training were completely absent in those who had received formal musical training as children, such that performance was matched with their monolingual counterparts. The results indicate that early musical training has the potential to be an effective long-term intervention for individuals who experience weakness in phonological processing.

Keywords

Bilingualism, cognition, comprehension, language processing, music, musical training, phonological processing, spelling, phoneme

Submission date: 8 October 2019; Acceptance date: 23 March 2020

Introduction

Engagement in musical training has been shown to provide strong benefits for numerous areas of language processing, especially in the domain of phonology (Gordon et al., 2015; Herrera et al., 2011; Patscheke et al., 2016). The ability to analyze and manipulate language down to the smallest unit of sound (i.e., the phoneme) is essential for generating a meaningful interpretation of the speech signal. The phoneme provides the lowest level at which the acoustic signal can be converted into its auditory equivalent in order to allow access to the mentally stored representation of a spoken word. Moreover, being sublexical, phonemic processing allows the mental representation of a novel spoken word (or non-word) to be formed. Therefore, differences in the ability of a person to analyze and manipulate phonemes is likely to have an impact on their ability to process spoken words, especially when the words are novel, and there is evidence of this in bilinguals, even when

their second language (i.e., L2) was learnt early in life and has become their dominant language in adulthood (e.g., Nguyen-Hoan & Taft, 2010). Given the known benefits of musical training, then, the aim of the present study is to establish whether early musical training is associated with phonological performance for such “early L2-dominant bilinguals” that is at the same level as their monolingual counterparts. It has been previously shown that the phonological proficiency of bilingual children can be improved through musical training (e.g., Patscheke et al., 2016), but we do not know whether the level of

School of Psychology, The University of New South Wales, Sydney, NSW, Australia

Corresponding author:

Angela Liang, School of Psychology, University of New South Wales, Sydney, NSW 2052, Australia.

Email: angela.liang@unsw.edu.au



Creative Commons Non Commercial CC BY-NC: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 License (<https://creativecommons.org/licenses/by-nc/4.0/>) which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is attributed as specified

on the SAGE and Open Access pages (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

improvement through to adulthood is such that their speech processing is comparable to that of monolinguals.

Music-to-Language Transfer Effects in L1

In recent years, a growing body of literature has linked early musical training (i.e., learning an instrument) to enhancements in multiple areas of language processing such as speech and auditory perception (e.g., François & Schön, 2011; Tierney et al., 2015) and phonological processing skills (e.g., Gordon et al., 2015; Herrera et al., 2011; Patscheke et al., 2016; Seither-Preisler et al., 2014). For example, Magne et al. (2006) found that musically trained 7- to 9-year-olds were better able to detect small pitch inconsistencies in sentences than those who were not musically trained. Similarly, Chobert et al. (2011) demonstrated that musically trained 9-year-olds had enhanced sensitivity and increased brain responses to deviations in the duration and frequency of syllables, voice-onset-time, and speech segmentation compared to those without training. Considering that speech and music both rely heavily on auditory learning, it is unsurprising that both practices share commonalities at an acoustic level, such as neuronal substrates for auditory perception (e.g., Peretz et al., 2015), and utilize similar information processing systems and strategies (Kraus et al., 2009; Patel, 2011, 2014; Schön et al., 2010). Specifically, both music and speech rely on similar acoustic cues, including duration, frequency, pitch, and timbre (e.g., Kraus et al., 2009; Kraus & Chandrasekaran, 2010; Seither-Preisler et al., 2014). As such, music-to-language transfer effects are thought to arise from the mechanisms that musical training and speech processing share for learning sound categories (McMullen & Saffran, 2004; Patel, 2008; Patscheke et al., 2018).

Especially during early childhood where music and language abilities develop rapidly, the skills required to process minute acoustic variations in music can transfer to linguistic-related auditory processing tasks, namely, phonological processing (e.g., Anvari et al., 2002). Such music-to-phonological transfer effects stem from the similarities that both practices have in segmenting the relevant acoustic cues into smaller sound units. For example, phonological processing requires deconstruction of speech and identification of the individual sound units across variation in pitch, tempo, speaker, and context. Likewise, musical perception involves the segmentation of phrases and chords into notes. Further, the enhanced timing precision and metrical skills (i.e., rhythm in music) that is developed as a result of musical training, facilitates one's ability to process rhymes, segment words, and manipulate phonemes (e.g., Anvari et al., 2002; Huss et al., 2011; Tierney & Kraus, 2014; Woodruff Carr et al., 2014). The resulting enhancement in segmentation and auditory-related abilities has been repeatedly shown to enrich phonological processing in children of all ages from 4 to 9 years

(e.g., Anvari et al., 2002; Degé & Schwarzer, 2011; Gromko, 2005) as well as in adolescents (e.g., Tierney et al., 2015). Remaining long-term effects have also been found in adults who reported engaging in early musical training and regardless of whether their training ceased after several years (e.g., Dittinger et al., 2016; Zuk et al., 2013).

Phonological Processing in L2

Given that musical training has been shown to enhance phonological performance in the general population of adults, the question can be raised as to whether early L2-dominant bilingual adults, who experience phonological processing disadvantages, can also benefit from musical training. Specifically, Nguyen-Hoan and Taft (2010) found that early L2-dominant bilingual adults process units of sound at a larger grain-size than monolinguals, displaying less sensitivity to the phonemic level of processing. This was true despite the heterogeneity of L1 amongst the bilingual participants and the notion that L1-to-L2 transfer effects may vary based on the structural distance between the two languages (e.g., Holm & Dodd, 1996; Wang et al., 2003). Grain-sizes larger than the phoneme include the syllable as well as subsyllabic phoneme combinations such as complex onsets (i.e., when there is more than one consonant preceding the vowel, e.g., the /gr/ of the non-word /grɛŋk/, i.e., "grenk") and rimes (i.e., the vowel plus any subsequent consonants, e.g., the /ɛŋk/ in /grɛŋk/). Processing on the basis of such units results in the use of sub-optimal strategies for deciphering novel utterances and interpreting spoken discourse involving novel words for meaning. Such tasks require analysis at the sublexical level, in contrast to real words which can be recognized using any grain-size of processing, including the word as a whole. Thus, the non-native-like performance of early L2-dominant bilinguals in tasks that use novel words (Nguyen-Hoan & Taft, 2010) suggests that their auditory processing mechanisms have not been optimally attuned, which is a putative consequence of the delay in L2 acquisition compared to the time of first exposure to an L1 (see Gordon et al., 2015, for a review).

The Relationship between Musical Training and L2 Phonological Processing

Whilst a large body of literature supports the benefits of musical training on native language literacy and skills, less is known about its effects on L2 acquisition and performance. Given that phonological processing is a skill that underlies language performance in general, it might be expected that any advantages in phonological processing gained from musical training would not only be observed in an individual's native language, but in L2 processing as well. This was indeed shown to be true by Slevc and Miyake (2006) who found in Japanese-English late bilinguals (i.e., with no L2 immersion before the age of 12

years) that musical training and experience enhanced L2 receptive and productive phonology, as distinct from other areas of L2 proficiency such as syntax and lexical knowledge. Their findings suggest that engagement in musical training may facilitate the development of L2 sound mechanisms. Further, Herrera et al. (2011) found that preschool children, who were randomly allocated to engage in a musical training programme, demonstrated enhanced phonological awareness when processing Spanish regardless of whether it was their native or second language, and compared to those who did not receive training. Such an enhancement in phonological skills has also been observed in L1-dominant bilingual children with a variety of L1 and L2 language backgrounds such as Kurdish, Asian, and European languages (Patscheke et al., 2016). Together, these results not only substantiate the notion that musical training is an effective means of improving phonological performance, but also are fundamental in highlighting that its transfer effects are not limited to native or specific languages.

While the focus of the existing literature has been to explore the effects of musical training on bilingual children, the outcomes for early L2-dominant bilingual adults has not been examined. Whether early musical training enables phonological performance in these bilinguals to match that of monolinguals is currently unknown. While it is unlikely that late bilingual adults would be ever able to achieve native-like proficiency in L2, the subtleties of the processing differences exhibited by early L2-dominant bilinguals suggest that their phonological performance has the potential to match that of their monolingual peers when musically trained. If this is found to be the case, the resulting association between musical training and phonological processing may indicate training as a potentially effective non-targeted long-term intervention for this group of individuals.

Given that biological evidence for long-term training-driven neural plasticity in the auditory system has been found (e.g., White-Schwoch et al., 2013), long-term benefits of training on phonological processing in early L2-dominant bilinguals seems likely. For example, in a cross-sectional study, Skoe and Kraus (2012) found that adults who had received formal musical instruction during childhood had more robust brainstem responses to sound than those who had never received training, suggesting that neural changes associated with musical training during childhood persist into adulthood. Further, the musician's brain has been shown to have the ability to coordinate complex communication between various brain regions during different developmental stages, allowing an enhanced response to a broader range of auditory input relative to non-musicians (e.g., Besson et al., 2016; Carpentier et al., 2016; Patel, 2011). Musical training also activates and fine-tunes the auditory processing system and acoustic-related neural mechanisms across the lifespan (Patel, 2011, 2014), which are similarly utilized during

language acquisition and use. Therefore, early musical training, which begins during the period where a second language is still being acquired, may enhance phonological processing in both L1 and L2, and persist long-term.

Thus, the present study aimed to examine whether the phonological processing of early L2-dominant bilingual adults might be more native-like if they have had early musical training (i.e., beginning before the age of around 8 years) than was observed by Nguyen-Hoan and Taft (2010). Indeed, their performance might be indistinguishable from that of monolinguals. The ideal design to examine this would be to randomly assign participants to musical training and no musical training conditions in their childhood and then compare their phonological performance as adults. However, such a longitudinal design is obviously impracticable and, therefore, the present study uses a quasi-experimental approach by making use of individuals who have already received early musical training. While this does raise the issue of causality, which will be addressed at a later point, such a design at least allows for an examination of whether early musical training is associated with enhanced language processing. A factorial design was therefore employed, with the factors being language background (i.e., monolingual or bilingual) and musical training experience (i.e., with or without early musical training).

The same series of tasks examined by Nguyen-Hoan and Taft (2010) was adopted here: phoneme deletion, spelling-to-dictation, and auditory comprehension. In addition to these phonological tasks, visual memory was measured to determine whether any impact of musical training was restricted to phonological processing, or whether it generalized to other domains. Previous evidence suggests that visual memory is influenced by neither musical training (Ho et al., 2003) nor language background (Namazi & Thordardottir, 2010) and, therefore, it was expected that there would be no differences in visual memory performance between the groups.

Performance was compared between musically trained English monolinguals, musically trained English-L2 bilinguals, English monolinguals with no musical training, and English L2-bilinguals with no musical training. All participants had their formal education in an English-speaking country (primarily Australia) and hence were dominant in English while still being fluent in their L1. Although Nguyen-Hoan and Taft (2010) found differences in performance between bilinguals from different L1 backgrounds (e.g., Chinese versus non-morphosyllabic languages), the bilinguals as a group nevertheless exhibited poorer performance on the phonological processing tasks relative to the monolingual English speakers. Therefore, no attempt was made to test bilinguals from only one specific L1 language background, though about half of them spoke Chinese as their L1 (either Mandarin or Cantonese).

Table 1. Participant characteristics (SD in parentheses).

	No musical training		Musical training	
	Monolingual	Bilingual ^a	Monolingual	Bilingual ^a
Gender distribution (f = female, m = male)	20f, 10m	15f, 15m	22f, 8m	15f, 15m
Mean age (years)	20.8 (6.94)	22.1 (4.32)	19.2 (1.50)	19.6 (3.20)
Mean estimated age of L1 onset (years)	–	1.42 (1.20)	–	1.05 (0.99)
Mean estimated age of L2 English onset (years)	–	3.75 (1.22)	–	4.00 (1.81)
Mean estimated exposure to English in the past year (%)	–	78.36 (12.91)	–	77.67 (7.51)
Mean estimated age of commencing musical training (years)	–	–	6.23 (1.25)	5.38 (1.34)
Mean estimated duration of musical training (years)	–	–	8.73 (3.59)	9.42 (3.59)
Mean estimated frequency of musical engagement (times/week)	–	–	4.30 (2.01)	5.30 (2.02)
Mean estimated musical proficiency (7 = extremely proficient)	–	–	4.77 (1.48)	5.13 (0.94)

^aNote that 13 bilingual participants without musical training and 11 bilingual participants with training reported learning a third language in high school to elementary proficiency.

Method

Description of Participants

Participants were recruited on the basis of their responses to a pre-screening questionnaire that inquired about language background and musical training. All bilingual participants reported that they had learnt their non-English language prior to learning English, which had then become their dominant language. They also had completed all their formal education in an English-speaking country (primarily Australia) and reported native-like proficiency in English. Musically trained participants had received at least five years of formal training (i.e., learning an instrument) that commenced before the age of 8 years.¹ None of the participants reported having any problems with speech, hearing, or language development, and all were current students at the University of New South Wales (UNSW Sydney). There were four groups of participants, with $n = 30$ in each.

Detailed participant characteristics were ascertained from a questionnaire administered during the study and these are presented in Table 1. On average, the musical bilingual group commenced their musical training at an earlier age than did the musical monolingual group, $t(58) = 2.54$, $p = .014$, and there was a trend for bilinguals to have more frequent musical engagement per week than monolinguals, $t(58) = 1.77$, $p = .081$. Therefore, if musically trained monolinguals were to show an advantage over musically trained bilinguals on any of the performance measures to be outlined later, this cannot be attributed to the former having had more exposure to their training than

the latter. All other variables were matched between the relevant groups, all p 's $> .10$. As there was an uneven distribution of females and males across the four groups, however, sex was entered as a covariate in all analyses of the performance data in order to hold it constant as a factor.

Materials and Procedure

In addition to the three phonological tasks (adapted from Nguyen-Hoan & Taft, 2010) and the visual memory task, non-verbal intelligence was measured using Raven's Advanced Progressive Matrices (RAPM: Bors & Stokes, 1998). All tasks were completed within a single one-hour session and are described as follows in the order in which they were administered.

Non-verbal Intelligence: RAPM. In order to ensure that any observed effects of language background or musical training on phonological processing could not be ascribed simply to differences in non-verbal intelligence, participants were given the shortened version of the RAPM. For each of the 12 items, a grid of patterns was presented consisting of eight figures with a ninth missing, and participants were asked to select from eight additional figures which one most appropriately completed the pattern. They were allowed to work through the items within a 10-minute time limit, and performance was measured by the number of correct responses out of 12. If the groups differed on their RAPM scores, these were to be entered into the analyses as a covariate and, hence, held constant.

Phoneme Deletion. The phoneme deletion task involved 12 monosyllabic non-words comprising an onset and rime, orally presented to participants through headphones. Each non-word was constructed with a complex onset that included more than one consonant (e.g., /frouʃ/, i.e., “froash”, or /crɔ:m/, i.e., “crawm”). Stimuli were recorded by a female native speaker of English. After each presentation, subjects were instructed to say aloud each item without its initial sound, with 3 seconds to respond. If a single phoneme were deleted (e.g., deleting /f/ from /frouʃ/), then the phoneme removed would always be a consonant, producing a non-word on deletion (e.g., /rouʃ/). Participants were not told this in the instructions and were not given any practice items. All responses were digitally recorded and analyzed during each session. Performance was measured in terms of the number of phoneme-sized deletions made, with a greater number of phoneme-sized deletions indicating greater sensitivity to the phoneme.

Spelling-to-dictation. Fifteen monosyllabic non-words that were spoken and pre-recorded by the same female native English speaker as for the phoneme deletion task were presented to each participant via headphones. Each non-word was potentially confusable with at least one real word because of their phonetic similarity, for example, /fi:g/ (i.e., “feeg”) as opposed to /fi:g/ (i.e., “fig”). Subjects were instructed to write down the correct spelling on an answer sheet provided, and were allowed to work through the items at their own pace. Each item was presented only once. Performance was measured in terms of whether the given spelling could be read with the same pronunciation as the spoken non-word. For instance, *fime* and *fyme* were both accepted as a correct spelling of the item /fɪɪm/, but *thime* and *fimm* were not.

Auditory Comprehension. Through headphones, participants listened to 10 short passages containing novel proper names, and were then asked one open-ended comprehension question per passage. The passages were one to two sentences in length, for example: “As Cydric Canusto listened to a mandolin being played skilfully by the young girl in the inn, his thoughts travelled back several months and several hundred leagues to the northern town of Dargon.” The corresponding question being: “What instrument was being played in the inn?” Subjects were given only one opportunity to listen to each passage and question, and were instructed to write the most appropriate answer on the response sheet provided. Passages were pre-recorded by the same female native English speaker as in the other tasks. Performance was measured in terms of the number of correct responses to the comprehension questions.

Visual Memory Test. Participants were instructed to remember a series of 16 different shapes (e.g., a square or a cross), each filled with one of four possible colours (red, blue, yellow, or green). Each shape was presented for 5 seconds. The same

shapes were then immediately presented individually in black and white in a different order, and participants were asked to recall the correct colour for each shape by selecting it from a choice of the four possible colours. Participants were instructed to write the correct answer on the provided response sheet, and were able to complete the task in their own time. Visual memory performance was measured in terms of the number of correct responses on this task.

Results and Discussion

A two-way analysis of covariance (ANCOVA) was carried out for each of the aforementioned tasks, with language background and musical training as the two between factors. Mean scores across the four groups were compared, with sex entered as a covariate for each analysis.

Non-verbal Intelligence Test (RAPM)

Collapsed across language groups, musical training was associated with higher average scores on the RAPM compared to those without musical training, $F(1, 114) = 16.12$, $p < .001$ (see Figure 1a). No main effect of language background, $F(1, 114) = 1.38$, $p = .243$, nor interaction between the two factors, $F(1, 114) = .06$, $p = .803$, were found. While these results show a relationship between musical training and non-verbal intelligence, the direction of causality is unclear (see, e.g., Forgeard et al., 2008; Franklin et al., 2008). It is possible that musical training enhances non-verbal intelligence, but also that parents of children who demonstrate higher non-verbal aptitude are more likely to encourage engagement in early musical training. Regardless of the direction of causality, however, it is important to ensure that any differences between groups in terms of phonological processing difficulties cannot be ascribed to disparities in non-verbal intelligence. Therefore, RAPM scores were entered in all further analyses as a covariate (along with sex).

Phoneme Deletion

The occasions where a single phoneme was deleted from the presented non-words will be referred to as “phoneme-sized deletions”, and the mean percentage of these across the four groups can be seen in Figure 1b. The analysis found that across language groups, those with musical training made significantly more phoneme-sized deletions than those without musical training, $F(1, 114) = 18.68$, $p < .001$. Language background was not associated with performance on this task, $F(1, 114) = .05$, $p = .817$, nor was an interaction found between the two factors, $F(1, 114) = .63$, $p = .428$.

Collapsed across language background, then, those who were musically trained were more likely than non-musicians to interpret the first sound of an utterance as a single phoneme. For example, when asked to remove the first sound of the non-word /blæŋ/ (i.e., “blang”), musicians

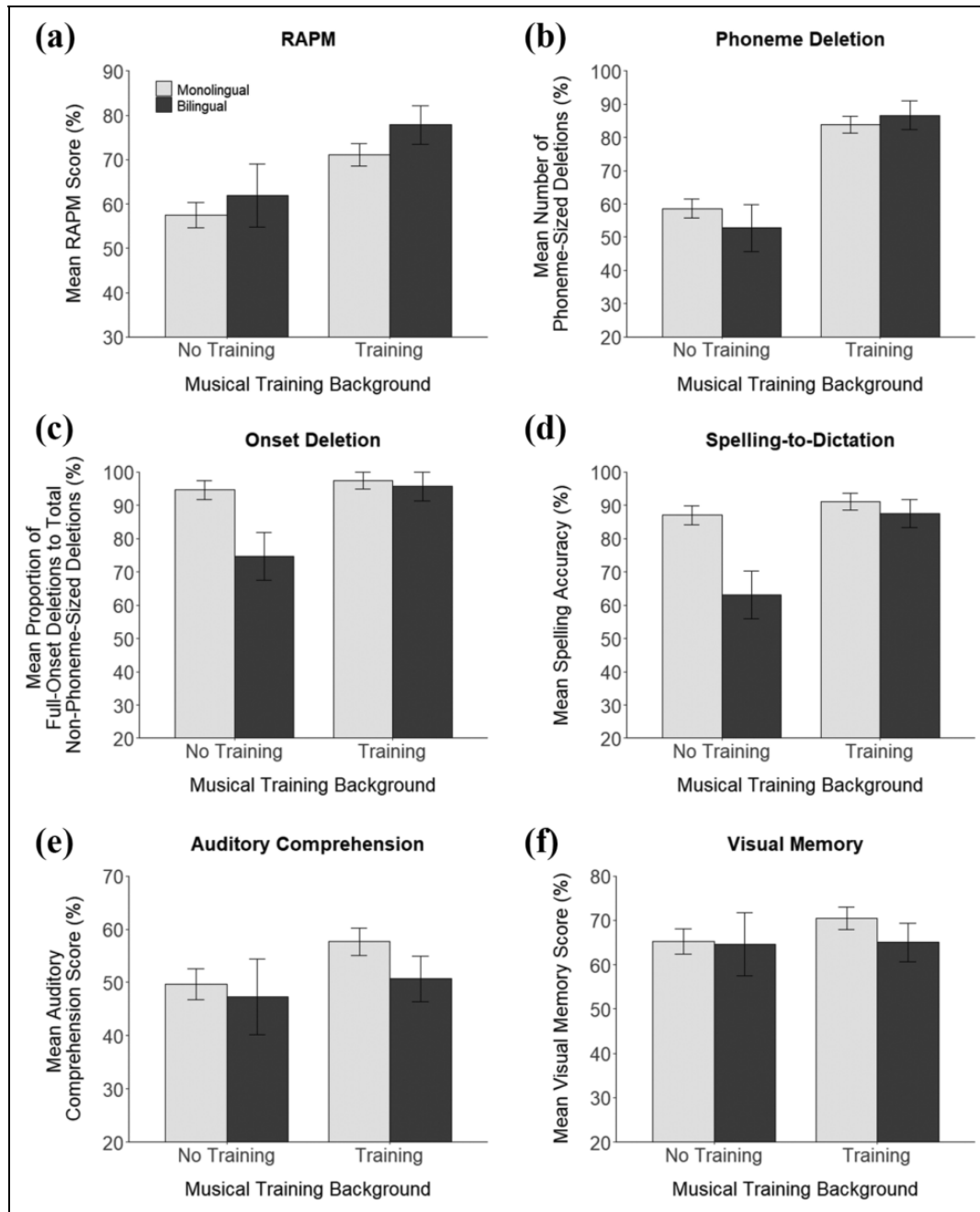


Figure 1. Mean percentage scores across the four groups for (a) the RAPM, (b) the phoneme deletion task in terms of phoneme-sized deletions, (c) the phoneme deletion task in terms of the proportion of Full-Onset Deletions out of all the non-phoneme-sized deletions, (d) the spelling-to-dictation task, (e) the auditory comprehension task, and (f) the visual memory task. The legend at the top left of Figure 1a applies to all plots. Error bars indicate standard error.

were more likely than non-musicians to produce /æŋ/ (i.e., by removing the phoneme /b/), with non-musicians having a higher tendency to produce /æŋ/ or /ŋ/, thus processing units of sound at a larger grain-size (i.e., by removing /bl/ or /blæ/). These findings suggest that musical training is associated with enhanced sensitivity to the phoneme, hence supporting a growing body of literature that highlights a positive relationship between musical training and improved phonemic awareness (e.g., Gromko, 2005;

Patscheke et al., 2018). Specifically, these findings support the notion that the similarities in the execution of musical and phonological skills with regard to segmenting relevant acoustic cues into smaller sound units facilitates one's ability to deconstruct and process words at the level of the phoneme (e.g., Anvari et al., 2002; Woodruff Carr et al., 2014). Further, children's awareness of rhyme develops before their awareness of phonemes, and rhyme proficiency is a strong predictor of later phonemic sensitivity (Carroll

et al., 2003). As robust associations between music-related rhythmic skills and rhyme awareness have been reported (e.g., Gordon et al., 2015; Huss et al., 2011), engagement in musical training during the period of language development is likely to benefit the rhyming system and, consequently, facilitate increased sensitivity to the phoneme.

While the impact of musical training was evident across language background, bilinguals and monolinguals displayed comparable performance in phoneme deletion. These results are inconsistent with previous findings that bilinguals process units of sound with larger grain-sized units than monolinguals (Nguyen-Hoan & Taft, 2010). However, a more in-depth analysis of the responses does provide evidence that language background was influential in the task after all. The deletions that were not phoneme-sized can be classified into two categories: Full-Onset Deletions (i.e., removal of an onset larger than a single phoneme, e.g. /fr/ deleted from /frouʃ/) and Other Responses (i.e., either removal of more than the onset, e.g. /frou/ deleted from /frouʃ/, or production of an erroneous phoneme, e.g. /ʌʃ/ in response to /frouʃ/). Full-Onset Deletions indicate greater sensitivity to phonological structure than do Other Responses inasmuch as the former at least implies a subsyllabic analysis into onsets and rimes, even if not an analysis into individual phonemes. Figure 1c illustrates the proportion of non-phoneme-sized deletions that were Full-Onset Deletions for each of the four groups. A two-way ANCOVA revealed a significant main effect of language, $F(1, 57) = 4.80, p < .05$, with monolinguals making a larger proportion of Full-Onset Deletions amongst their non-phoneme-sized responses than did bilinguals (and hence fewer Other Responses). Importantly, the analysis also revealed a significant interaction between language background and musical training, $F(1, 57) = 4.08, p < .05$, with a simple-effects analysis showing that of those without musical training, bilinguals were less likely than monolinguals to make a Full-Onset Deletion as opposed to an Other Response, $t(57) = 3.20, p < .01$, but that this effect was not found amongst the musically trained, $t(57) = 0.13, p > .10$.

Thus, although there were no quantitative differences in non-phoneme-sized deletions made between monolinguals and bilinguals, significant disparities between the types of such deletions were present. Given the non-word item /stɒnd/ (i.e., “stond”), for example, removal of the initial phoneme /s/ (i.e., making a phoneme-sized deletion, and hence producing /tɒnd/) displays sensitivity at the phoneme level and the ability to deconstruct the complex onset (i.e., /st/) into its smaller phonological units (i.e., /s/ and /t/). Removal of the full complex onset, /st/ (i.e., making a Full-Onset Deletion), demonstrates that the individual interprets the onset as an indivisible unit, hence displaying decreased sensitivity to the phoneme. However, producing a Full-Onset Deletion does indicate sensitivity to at least the onset-rime structure of which monosyllabic words are comprised (e.g., Treiman, 1986). On the other hand, the

most inappropriate category of non-phoneme-sized deletions, Other Responses, involves the removal of more than a complex onset, for instance /stɒ/ or /stɒn/ being deleted from /stɒnd/, or producing an incorrect response such as /dɒn/. Making an Other Response highlights a lack of sensitivity to any linguistically defined internal word structure, not just phonemic structure.

Of those who produced non-phoneme-sized deletions, bilinguals without musical training showed less sensitivity to the onset-rime structure of words than those with training, demonstrating the use of sub-optimal strategies when processing spoken English at the sublexical level compared to native monolingual speakers. However, the bilingual insensitivity to internal word structure was completely eliminated in those who had received early musical training. These findings are consistent with the notion that musical training can facilitate the attention paid to sounds and their manipulation (Ahissar et al., 2009; Patel, 2011), and therefore, has a positive influence on how speech sounds are analyzed in L2 (e.g., Anvari et al., 2002; Besson et al., 2011; McMullen & Saffran, 2004; Patscheke et al., 2018). Moreover, the fact that bilinguals with musical training exhibited enhanced sensitivity to rimes compared to those without training is consistent with the reported association between music-related rhythmic skills and rhyme awareness (e.g., Huss et al., 2011). The present results are remarkable in demonstrating that even the most subtle processing differences exhibited by early L2-dominant bilinguals are absent in those with musical training and persist in the long-term. These findings provide a preliminary indication that musical training may be a potentially effective tool in improving areas of phonological weakness at a low level of analysis in L2.

Spelling-to-dictation

Across musical training background, there was a significant main effect of language, $F(1, 114) = 21.06, p < .001$, with monolinguals having higher spelling accuracy of non-words than bilinguals (see Figure 1d). In addition, those with musical training performed better than those without, $F(1, 114) = 12.47, p = .001$, and there was a significant interaction between the two factors, $F(1, 114) = 10.05, p < .01$. An analysis of simple effects revealed that the monolinguals without musical training showed an advantage over bilinguals without musical training, $t(114) = 5.53, p < .001$, while there was no difference between the musically trained monolingual and bilingual groups, $t(114) = 1.07, p > .10$. These results indicate that early L2-dominant bilinguals do not attain native-like proficiency in their phonological analysis of spoken utterances unless they were immersed in musical training at a young age.

The results from this task suggest that the difficulty bilinguals experience in spelling non-words arises either from an inappropriate processing of the phonological form, which is then reflected in an inaccurate spelling, or from

difficulty in converting sublexical phonological information into its orthographic form. A further analysis of the error responses is more consistent with the former. In particular, it can be argued that if people have difficulty processing the phonological form of an utterance, they may well assimilate it to a similar phonological form that they already know. As such, they would be more likely to give a real word as their response than someone who processed the speech signal correctly and hence realized that a real word was an inappropriate response. On analyzing the proportion of real word errors made amongst the total number of errors, an interaction was again found between language background and musical training, $F(1, 56) = 4.74, p < .05$, with bilinguals being more likely than monolinguals to make real word errors when they had no musical training, $t(56) = 3.59, p = .001$, but not when they did, $t(56) = 0.12, p > .10$. These results suggest that rather than there being a difficulty for bilinguals without training to convert spoken utterances into their orthographic form, their inaccuracy in spelling arises from a difficulty in processing the speech signal, such that they resort to reporting a similar word that is already stored. As this difficulty was not present in those with musical training, the findings indicate that engagement in early musical training is positively associated with enhanced processing of the speech signal in a dominant L2.

Auditory Comprehension

Although Figure 1e indicates that monolinguals performed better than bilinguals in the comprehension task, this did not reach statistical significance, only being a strong trend, $F(1, 114) = 3.29, p = .072$. Further, no impact of musical training was observed either as a main effect, $F(1, 114) = 2.00, p > .10$, or as an interaction with language background, $F(1, 114) = .59, p > .10$. So, given that the bilinguals did not differ from the monolinguals even when musically untrained (contrary to the findings of Nguyen-Hoan & Taft, 2010), there was no disadvantage for musically trained bilinguals to overcome.

It should be noted, however, that while Nguyen-Hoan and Taft (2010) observed a bilingual effect on this task regardless of the nature of L1, the Chinese bilinguals in that study nevertheless performed more poorly than the non-Chinese bilinguals. Such a finding suggests that bilinguals with an L1 that does not utilize subsyllabic processing (i.e., Chinese) have a particular difficulty in processing spoken English passages containing novel words, where subsyllabic processing is required. To examine this here, a post-hoc analysis separating the present sample of bilinguals with a Chinese and non-Chinese L1 was conducted in order to determine whether the type of L1 affected auditory comprehension, and a significant interaction between L1 and musical training background was indeed observed, $F(1, 54) = 4.69, p = .035$. A simple-effects analysis showed that, of those without musical training, Chinese bilinguals showed a disadvantage relative to non-Chinese bilinguals,

$t(54) = 2.45, p = .018$, but this effect was not found amongst the musically trained, $t(54) = .61, p > .10$. Therefore, the present results imply that at least bilinguals with a Chinese L1 utilize sub-optimal strategies when processing novel English speech for meaning, but that this processing disadvantage is eliminated in those with early musical training.

It should be noted, however, that the same post-hoc analysis conducted for the phoneme deletion and non-word spelling tasks failed to reveal any significant impact of having Chinese as L1. Given that Nguyen-Hoan and Taft (2010) did find a greater disadvantage for Chinese/English bilinguals than those with another L1 when performing the phoneme deletion task (though not when spelling non-words), the findings of the present analysis for the comprehension task should be noted with caution, and the differences in processing between bilinguals with a Chinese and non-Chinese L1 need confirmation through future research.

Visual Memory

Figure 1f shows the mean percentage correct visual memory responses across the four groups. In support of the existing literature (Ho et al., 2003; Namazi & Thordardottir, 2010), there were no main effects of either language background, $F(1, 114) = 2.11, p > .10$, or musical training, $F(1, 114) = .01, p > .10$, and neither was there an interaction between these two factors, $F(1, 114) = .87, p > .10$.

The fact that there was no effect of musical training on visual memory strengthens the notion that musical training is not associated with better performance on all cognitive abilities, with memory for colour being one such ability. This is not surprising since there is no obvious relationship between the cognitive mechanisms underlying memory for colour and musical performance.

General Discussion

The aim of the present study was to examine whether any phonological processing disadvantages that are found to exist in early L2-dominant bilingual adults are not observed in those who have engaged in early musical training. This is indeed what was demonstrated in the present study, which might suggest that commencing such training at an age where the auditory processing system is still maturing facilitates the development of the phonological system. Specifically, those with training showed increased onset-rime sensitivity and more accurate spelling of non-words, indicating an enhanced capacity to process the internal phonological structure of utterances. Moreover, Chinese bilinguals with early musical training did not exhibit the difficulties in processing English speech for meaning than seen in those without training, hence demonstrating a benefit that extends to everyday language functioning. The present findings are consistent with previous research

revealing music-to-language transfer effects (e.g., François & Schön, 2011; Gordon et al., 2015; Herrera et al., 2011; Patscheke et al., 2016) and the notion that the training-driven enhancements of the networks that speech and music processing share are particularly beneficial for bilinguals who experience phonological difficulties (Patel, 2011, 2014).

While the focus of the study was not to determine how, but whether there was an association between early musical training and the elimination of phonological processing difficulties, the results demonstrate that the processing difficulties that early L2-dominant bilinguals experience are not so intractable for musical training to overcome. Given that early musical training has previously been proposed as a remediation strategy for individuals with noticeable phonological difficulties such as those with spoken language impairment or a reading disability (Bryant & Bradley, 1985), the findings from the present study suggest that early musical training may, too, potentially be an effective intervention for bilinguals who experience phonological difficulties that are so subtle that they tend to go unnoticed.

Having said this, however, it is important to note that such a promising conclusion can only be seen as suggestive. This is because, as with any quasi-experimental design, the results are subject to influence by confounding variables and, therefore, the causality of the relationship between musical training and language performance remains an open question. That is, it cannot be established from the present study whether musical training directly enhances the phonological system or whether it is those who have a more developed phonological system that are encouraged to engage in musical training. Nevertheless, the outcomes of the present study support a growing number of cross-sectional findings indicating a strong relationship between early musical training and enhanced phonological performance in adulthood (e.g., Dittinger et al., 2016; Zuk et al., 2013). Given that both musical and phonological processing require similar use of the auditory system (e.g., Kraus et al., 2009; Peretz et al., 2015; Schön et al., 2010), it is possible that individuals who display stronger phonological performance adapt more readily to or excel in their musical training and, hence, are more inclined to pursue it. Moreover, there may be underlying factors that are shared by and influence both musical training and L2 performance. For example, parents of children who undertake or maintain musical practice may be more engaged in their child's education, thus providing more enriched environments for optimal L2 acquisition, than parents who do not enrol their child in musical training. It is possible that, along with music classes, these children were also immersed in L2 development programmes or received targeted linguistic interventions, which subsequently enriched their phonological aptitude. Exposure to both musical training and linguistic interventions might also be determined by socio-economic circumstances, and while this was not

established in the present research, is worthy of future exploration.

Thus, it is possible that a more sophisticated relationship between music and phonological skills exists than the simple idea that it is musical training that directly enhances phonological abilities. Whilst the current design was limited in the number of external variables that could be considered, the factors that were accounted for with regard to language and musical training history provide a strong platform for future research.

Contributorship

AL researched the literature, conceived and conducted the study, analyzed the data, and wrote the first draft of the manuscript. MT was involved in the study design, interpretation of data, critical review of the manuscript, and supported the study throughout. Both authors reviewed and edited the manuscript and approved the final version of the manuscript.


Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Angela Liang  <https://orcid.org/0000-0002-3321-385X>

Action Editor

Daniela Sammler, Max Planck Institute for Human Cognitive and Brain Sciences.

Peer Review

Sabrina Turker, Max-Planck-Institut für Kognitions- und Neurowissenschaften, Lise Meitner Research Group "Cognition and Plasticity".

One anonymous reviewer.

Note

1. "Early musical training" is defined in this study as commencing before the age of 8 years based on a consideration of the realistic age at which most children start learning a musical instrument and existing evidence that the phonological system is still developing at this age (e.g., Vogel & Raimy, 2002).

References

- Ahissar, M., Nahum, M., Nelken, I., & Hochstein, S. (2009). Reverse hierarchies and sensory learning. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 364 (1515), 285–299. <http://doi.org/10.1098/rstb.2008.0253>
- Anvari, S. H., Trainor, L. J., Woodside, J., & Levy, B. A. (2002). Relations among musical skills, phonological processing, and early reading ability in preschool children. *Journal of Experimental Child Psychology*, 83(2), 111–130. [http://doi.org/10.1016/S0022-0965\(02\)00124-8](http://doi.org/10.1016/S0022-0965(02)00124-8)

- Besson, M., Chobert, J., & Marie, C. (2011). Transfer of training between music and speech: Common processing, attention, and memory. *Frontiers in Psychology*, 2, 94. <http://doi.org/10.3389/fpsyg.2011.00094>
- Bors, D. A., & Stokes, T. L. (1998). Raven's advanced progressive matrices: Norms for first-year university students and the development of a short form. *Educational and Psychological Measurement*, 58(3), 382–398. <http://doi.org/10.1177/0013164498058003002>
- Bryant, P. E., & Bradley, L. (1985). *Children's reading problems*. Basil Blackwell.
- Carpentier, S. M., Moreno, S., & McIntosh, A. R. (2016). Short-term music training enhances complex, distributed neural communication during music and linguistic tasks. *Journal of Cognitive Neuroscience*, 28(10), 1603–1612. http://doi.org/10.1162/jocn_a_00988
- Carroll, J. M., Snowling, M. J., Stevenson, J., & Hulme, C. (2003). The development of phonological awareness in preschool children. *Developmental Psychology*, 39(5), 913. <http://doi.org/10.1037/0012-1649.39.5.913>
- Chobert, J., Marie, C., François, C., Schön, D., & Besson, M. (2011). Enhanced passive and active processing of syllables in musician children. *Journal of Cognitive Neuroscience*, 23(12), 3874–3887. http://doi.org/10.1162/jocn_a_00088
- Degé, F., & Schwarzer, G. (2011). The effect of a music program on phonological awareness in preschoolers. *Frontiers in Psychology*, 2, 124. <http://doi.org/10.3389/fpsyg.2011.00124>
- Dittinger, E., Barbaroux, M., D'Imperio, M., Jäncke, L., Elmer, S., & Besson, M. (2016). Professional music training and novel word learning: From faster semantic encoding to longer-lasting word representations. *Journal of Cognitive Neuroscience*, 28(10), 1584–1602. http://doi.org/10.1162/jocn_a_00997
- Forgeard, M., Winner, E., Norton, A., & Schlaug, G. (2008). Practicing a musical instrument in childhood is associated with enhanced verbal ability and nonverbal reasoning. *PLOS One*, 3(10), e3566. <http://doi.org/10.1371/journal.pone.0003566>
- François, C., & Schön, D. (2011). Musical expertise boosts implicit learning of both musical and linguistic structures. *Cerebral Cortex*, 21(10), 2357–2365. <http://doi.org/10.1093/cercor/bhr022>
- Franklin, M. S., Sledge Moore, K., Yip, C. Y., Jonides, J., Rattray, K., & Moher, J. (2008). The effects of musical training on verbal memory. *Psychology of Music*, 36(3), 353–365. <http://doi.org/10.1177/0305735607086044>
- Gordon, R. L., Fehd, H. M., & McCandliss, B. D. (2015). Does music training enhance literacy skills? A meta-analysis. *Frontiers in Psychology*, 6, 1777. <http://doi.org/10.3389/fpsyg.2015.01777>
- Gromko, J. E. (2005). The effect of music instruction on phonemic awareness in beginning readers. *Journal of Research in Music Education*, 53(3), 199–209. <http://doi.org/10.1177/002242940505300302>
- Herrera, L., Lorenzo, O., Defior, S., Fernandez-Smith, G., & Costa-Giomi, E. (2011). Effects of phonological and musical training on the reading readiness of native-and foreign-Spanish-speaking children. *Psychology of Music*, 39(1), 68–81. <http://doi.org/10.1177/0305735610361995>
- Ho, Y. C., Cheung, M. C., & Chan, A. S. (2003). Music training improves verbal but not visual memory: Cross-sectional and longitudinal explorations in children. *Neuropsychology*, 17(3), 439. <http://doi.org/10.1037/0894-4105.17.3.439>
- Holm, A., & Dodd, B. (1996). The effect of first written language on the acquisition of English literacy. *Cognition*, 59, 119–147. [http://doi.org/10.1016/0010-0277\(95\)00691-5](http://doi.org/10.1016/0010-0277(95)00691-5)
- Huss, M., Verney, J. P., Fosker, T., Mead, N., & Goswami, U. (2011). Music, rhythm, rise time perception and developmental dyslexia: Perception of musical meter predicts reading and phonology. *Cortex*, 47(6), 674–689. <http://doi.org/10.1016/j.cortex.2010.07.010>
- Kraus, N., & Chandrasekaran, B. (2010). Music training for the development of auditory skills. *Nature Reviews Neuroscience*, 11(8), 599–605. <http://doi.org/10.1038/nrn2882>
- Kraus, N., Skoe, E., Parbery-Clark, A., & Ashley, R. (2009). Experience-induced malleability in neural encoding of pitch, timbre, and timing: Implications for language and music. *Annals of the New York Academy of Sciences*, 1169(1), 543–557. <http://doi.org/10.1111/j.1749-6632.2009.04549.x>
- Magne, C., Schön, D., & Besson, M. (2006). Musician children detect pitch violations in both music and language better than nonmusician children: Behavioral and electrophysiological approaches. *Journal of Cognitive Neuroscience*, 18(2), 199–211. <http://doi.org/10.1162/jocn.2006.18.2.199>
- McMullen, E., & Saffran, J. R. (2004). Music and language: A developmental comparison. *Music Perception: An Interdisciplinary Journal*, 21(3), 289–311. <http://doi.org/10.1525/mp.2004.21.3.289>
- Namazi, M., & Thordardottir, E. (2010). A working memory, not bilingual advantage, in controlled attention. *International Journal of Bilingual Education and Bilingualism*, 13(5), 597–616. <http://doi.org/10.1080/13670050.2010.488288>
- Nguyen-Hoan, M., & Taft, M. (2010). The impact of a subordinate L1 on L2 auditory processing in adult bilinguals. *Bilingualism: Language and Cognition*, 13(2), 217–230. <http://doi.org/10.1017/S1366728909990551>
- Patel, A. D. (2008). Science & music: Talk of the tone. *Nature*, 453(7196), 726. <http://doi.org/10.1038/453726a>
- Patel, A. D. (2011). Why would musical training benefit the neural encoding of speech? The OPERA hypothesis. *Frontiers in Psychology*, 2, 142. <http://doi.org/10.3389/fpsyg.2011.00142>
- Patel, A. D. (2014). Can nonlinguistic musical training change the way the brain processes speech? The expanded OPERA hypothesis. *Hearing Research*, 308, 98–108. <http://doi.org/10.1016/j.heares.2013.08.011>
- Patscheke, H., Degé, F., & Schwarzer, G. (2016). The effects of training in music and phonological skills on phonological awareness in 4-to 6-year-old children of immigrant families. *Frontiers in Psychology*, 7, 1647. <http://doi.org/10.3389/fpsyg.2016.01647>
- Patscheke, H., Degé, F., & Schwarzer, G. (2018). The effects of training in rhythm and pitch on phonological awareness in

- four-to six-year-old children. *Psychology of Music*, 47(3), 376–391. <http://doi.org/10.1177/0305735618756763>.
- Peretz, I., Vuvar, D., Lagrois, M. É., & Armony, J. L. (2015). Neural overlap in processing music and speech. *Philosophical Transactions of the Royal Society of London B: Biological Science*, 370, 20140090. <http://doi.org/10.1098/rstb.2014.0090>
- Schön, D., Gordon, R., Campagne, A., Magne, C., Astésano, C., Anton, J. L., & Besson, M. (2010). Similar cerebral networks in language, music and song perception. *Neuroimage*, 51(1), 450–461. <http://doi.org/10.1016/j.neuroimage.2010.02.023>
- Seither-Preisler, A., Parncutt, R., & Schneider, P. (2014). Size and synchronization of auditory cortex promotes musical, literacy, and attentional skills in children. *Journal of Neuroscience*, 34(33), 10937–10949. <http://doi.org/10.1523/JNEUROSCI.5315-13.2014>
- Skoe, E., & Kraus, N. (2012). A little goes a long way: How the adult brain is shaped by musical training in childhood. *The Journal of Neuroscience*, 32(34), 11507–11510. <http://doi.org/10.1523/JNEUROSCI.1949-12.2012>
- Slevc, L. R., & Miyake, A. (2006). Individual differences in second-language proficiency: Does musical ability matter?. *Psychological Science*, 17(8), 675–681. <http://doi.org/10.1111/j.1467-9280.2006.01765.x>
- Tierney, A., & Kraus, N. (2014). Auditory-motor entrainment and phonological skills: precise auditory timing hypothesis (PATH). *Frontiers in Human Neuroscience*, 8, 949. <http://doi.org/10.3389/fnhum.2014.00949>
- Tierney, A., Krizman, J., & Kraus, N. (2015). Music training alters the course of adolescent auditory development. *Proceedings of the National Academy of Sciences*, 112(32), 10062–10067. <http://doi.org/10.1073/pnas.1505114112>
- Treiman, R. (1986). The division between onsets and rimes in English syllables. *Journal of Memory and Language*, 25(4), 476–491. [http://doi.org/10.1016/0749-596X\(86\)90039-2](http://doi.org/10.1016/0749-596X(86)90039-2)
- Vogel, I., & Raimy, E. (2002). The acquisition of compound vs. phrasal stress: The role of prosodic constituents. *Journal of Child Language*, 29(2), 225–250. <http://doi.org/10.1017/S0305000902005020>
- Wang, M., Koda, K., & Perfetti, C. A. (2003). Alphabetic and nonalphabetic L1 effects in English word identification: A comparison of Korean and Chinese English L2 learners. *Cognition*, 87, 129–149. [http://doi.org/10.1016/S0010-0277\(02\)00232-9](http://doi.org/10.1016/S0010-0277(02)00232-9)
- White-Schwoch, T., Carr, K. W., Anderson, S., Strait, D. L., & Kraus, N. (2013). Older adults benefit from music training early in life: Biological evidence for long-term training-driven plasticity. *Journal of Neuroscience*, 33(45), 17667–17674. <http://doi.org/10.1523/JNEUROSCI.2560-13.2013>
- Woodruff Carr, K., White-Schwoch, T., Tierney, A. T., Strait, D. L., & Kraus, N. (2014). Beat synchronization predicts neural speech encoding and reading readiness in preschoolers. *Proceedings of the National Academy of Sciences*, 111, 14559–14564. <http://doi.org/10.1073/pnas.1406219111>
- Zuk, J., Ozernov-Palchik, O., Kim, H., Lakshminarayanan, K., Gabrieli, J. D., Tallal, P., & Gaab, N. (2013). Enhanced syllable discrimination thresholds in musicians. *PLOS One*, 8(12), e80546. <http://doi.org/10.1371/journal.pone.0080546>