Musical Aptitude and the Acquisition of Pitch and Duration in L2 Vowels

Rebecca Marianne Lund

Introduction

It has often been said that musicians have an advantage when learning new languages. This notion ties into the general discussion of the relationship between speech and music. Speech and music are both auditory signals unique to humans and are based on the same acoustic parameters (Chobert and Besson 2013). Furthermore, neurobiological studies have shown similarities in brain processing of music and speech stimuli, and four factors have been identified for music-related neuronal plasticity: years of training, amount of training, age at which training began, and musical aptitude (Kraus and Chandrasekaran 2010). Of these factors, this article will focus on musical aptitude, which is an inherent sensitivity towards aspects of musical sounds such as pitch and timing, which are also features of vowels (Kraus et al. 2009). The inherent sensitivity improves the individual's perception and discrimination of sounds and therefore increases their potential of becoming an accomplished musician (Talamini et al. 2018; Suzukida 2021). Evidence concerning the influence of musical aptitude on the acquisition of temporal and spectral features of vowels of an L2 will be reviewed in this article. Firstly, differences between using musical experience and musical aptitude as measures of musical ability will be examined. Secondly, neurobiological evidence of the effect of music on speech processing will be presented, and lastly, the mediation of the link between sensitivity to non-native speech-sound processing and musical aptitude will be explored.

Difference between musical aptitude and musical experience

Musical aptitude and musical experience, terms that will be elaborated on later, have been considered to constitute musical ability. In a study from 2006 by Slevc and Miyake it was examined whether musical ability could account for variance in L2 proficiency when measured against other known factors influencing L2 acquisition, e.g., age of arrival in L2 country and length of residence in that country. The participants were adult native speakers of Japanese, who moved to the United

Leviathan: Interdisciplinary Journal in English (ISSN: 2446-3981), No. 9, 2023.

States after the age of 11. Their language skills were assessed in receptive phonology, productive phonology, syntax, and lexical knowledge, and their musical ability was measured with three subtests of Wing Measures of Musical Talent test, which tests individuals' musical aptitude. The results showed that musical ability accounted for differences between participants in "L2 phonological ability (both receptive and productive) even when controlling for other factors but did not explain unique variance in L2 syntax or lexical knowledge" (Slevc and Miyake 2006, 675). Thus, musical ability was found to influence the phonological aspects of L2 acquisition. However, the study did not clearly state whether "musical ability" referred to musical aptitude or experience or both. Although the authors used an aptitude test to measure the participants' musical ability, they did not disclose the participants' experience with music. This raises the question of whether musical experience or aptitude is the most relevant music-related factor for L2 learning.

Many studies have used musical experience as a measure of musical ability. Experience is often measured using questionnaires, where participants are asked how many years of music training they have received. This method is not without weaknesses, as it raises questions of how to define "experience" and "musician", and how to quantify "experience". Firstly, years of training is an unreliable measure as two people who have spent an equal amount of years training, may differ in experience if one trained five times a week for 10 years and the other trained once a week. Secondly, one musician may be more gifted than another musician even though they have received the same amount of training (Nardo and Reiterer 2009, 215). Lastly, the musical experience approach fails to identify individuals who have not received any musical training but who are nevertheless musically gifted, as well as those who have received extensive training but remain unskilled (Law and Zentner 2012).

The question of how to define a musician arises when considering which linguistic feature is to be examined, as the relevance of different musicians' expertise may vary. For instance, the expertise of a percussionist and a violinist differs, as string instruments must be tuned, which requires a "greater sensitivity to small pitch changes … compared to other instruments that usually do not require such adjustments (keyboard or percussion)" (Gottfried 2007, 236). Furthermore, a subcategorization of musicians into "vocalists" and "instrumentalists" has been proposed, as singing is the musical ability most similar to language (Christiner and Reiterer 2016, 133), and as it has been found that although any musical training improves an individual's perceptual abilities, only "vocal motor training improves the ability to repeat unfamiliar speech … [and] to memorize new speech patterns more easily and faster" (Christiner and Reiterer 2016, 149-150).

One way to avoid the above-mentioned challenges is to use musical aptitude as a measure of musical ability instead of musical experience. Musical aptitude is defined as an inherent sensitivity towards aspects of musical sounds, e.g., pitch and timing, and can be measured in individuals with no musical experience. This sensitivity towards musical sounds improves the individual's perception and discrimination of sounds and increases their potential of becoming an accomplished musician (Talamini et al. 2018: Suzukida 2021). Musical aptitude, as a measure of ability, relies on aptitude tests. However, many tests exist and overlap in their subcomponents but differ slightly as "each test has been created from a different perspective [which has] generated a series of theoretical proposals and empirical approaches that do not always coincide" (Nardo and Reiterer 2009, 219).

Studies of music and language often compare the processing of musical sounds and vowels, as they share spectral and temporal features, including pitch and timing (Lidji et al. 2010). Pitch refers to the "organization of sound on an ordered scale (low versus high pitch) and is a subjective percept of the frequency of the sound" (Kraus and Chandrasekaran 2010), and corresponds in complex sounds, e.g., speech and musical harmonies, to the fundamental frequency (Kraus et al. 2009, 544). Timing refers to the temporal aspects of sounds, e.g., the duration of vowels or tones (Kraus et al. 2009, 544), and will now be referred to as duration. It has been hypothesized that an inherent sensitivity towards duration and pitch in musical sounds should improve the individual's perception and production of not only musical sounds but also of L2 vowels.

The relationship between L2 vowel acquisition and musical aptitude was examined in three studies by Milovanov et al. from 2008, 2009 and 2010. In the first study, the authors examined whether Finnish children with advanced English production skills had a higher musical aptitude, as measured by the Seashore aptitude test, than children with less-advanced production skills. The children had limited experience with the English phonemic system, as they were first- or second-year students of English (Milovanov et al. 2008). Two language tests were conducted, one where the participants' L2 production ability was tested by them repeating words containing phonemes that do not exist in Finnish, e.g., /θ/, /tʃ/, and /ʒ/ (Milovanov et al. 2008). The other language test was a phonemic discrimination test in which the participants had to "distinguish the phonemic dissimilarities between English and Finnish through triplets based on minimal pair contrasts" e.g., /d/ and /ð/ in the triplet they-day-they (Milovanov et al. 2008). The participants had an 8-week English pronunciation course and a post-test, similar to the discrimination test, which formed the basis for dividing the children into two groups, advanced pronunciation group (APG) and less-advanced pronunciation group (LPG). The musical experience varied similarly within each group and experience was therefore considered irrelevant to the study (Milovanov et al. 2008). Two musical tests were conducted. The first was a chord discrimination test similar to the phonemic discrimination test with the stimuli consisting of a C-major chord played on a violin with two deviants consisting of a mistuned major third being either 2% or 4% flat (Milovanov et al. 2008). The participants' musical aptitude was measured with the Seashore test and as shown in Table 1, the APG scored significantly higher, as indicated with three asterisks, in five of the subtests but the groups did not differ in the time and loudness subtests (Milovanov et al. 2008).

	APG (%)	LPG (%)
Pitch	72 +/- 6 ***	34 +/- 6
Timbre	70 +/- 6***	46 +/- 7
Rhythm	81 +/- 3***	61 +/- 6
Tonality	80 +/- 3***	60 +/- 6
Time	62 +/- 6	60 +/- 6
Loudness	67 +/- 6	59 +/- 6
General musical aptitude	72 +/- 3***	55 +/- 4

Table 1. The Seashore musical aptitude test scores measuring correct reports in percentages between the Advanced Pronunciation Group (APG) and the Less-advanced pronunciation group (LPG). Asterisks indicate significant mean differences between the groups with a significance level of $p < 0.001$.

This result corroborated the authors' proposal that "language skills, both in production and discrimination, are interconnected with perceptual musical skills" (Milovanov et al. 2008). The proposal was tested by the same authors in their 2010 study where Finnish adults, who were either English philology students, non-musicians, or choristers, were tested in the same Seashore subtests, as in the 2008 study, and on their English production skills. The choristers had the highest score in all Seashore subtests. The English philology students performed better than the non-musicians in five subtests, but they had similar scores in the pitch and rhythm subtests (Milovanov et al. 2010). In the production test, the choristers were as good as the philology students, while the non-musicians performed significantly worse than the other groups. A relation was therefore found for both the influence of language on musical aptitude and the influence of musical aptitude on language skills, as the philology students and the choristers scored similarly in both tests (Milovanov et al. 2010).

In their 2008 study Milovanov et al. were also interested in examining whether children with advanced L2 skills represented "musical sound features more readily in the pre-attentive level of neural processing compared to children with less-advanced production skills" (Milovanov et al. 2008). This was done with the use of the Mismatch Negativity paradigm.

Neurobiological evidence of the influence of music on speech processing

Mismatch negativity (MMN) is a part of event-related potentials (ERP) that are measured with the neuroimaging technique electroencephalography (EEG). The EEG measures the electrical activity of the brain via electrodes placed on the scalp. As the brain responds to multiple stimuli simultaneously, it is necessary to extract the electrical response caused by a specific stimulus from the overall EEG chart in order to examine the brain's response to this single stimulus. This extracted response is the event-related potential (Luck 2014, 4). While the ERP is a measure of the brain's response to a specific stimulus, the MMN is an "index of automatic change-detection which is elicited following any change to an established regularity in sensory stimuli, including sound" (Fitzgerald and Todd 2020). The MMN is often studied using an oddball discrimination task, where a series of the standard stimulus, often a sound, is presented with the deviant stimulus irregularly distributed within the series. If the standard sound has been neurally encoded in the participant then a spike will occur in the EEG when the deviant sound is presented, indicating the presence of the MMN. The terms "standard" and "deviant" do not "refer to individual tones necessarily, but rather the neural representations of a regularity and a violating event which can vary in complexity" (Fitzgerald and Todd 2020). The MMN measures the pre-attentive processing skills of humans and can therefore be "observed without conscious attention to the sound stream", which suggests that "sophisticated sensory discrimination processes are initiated at the pre-attentive level" (Fitzgerald and Todd 2020). The MMN response can be influenced by the individual's linguistic and musical expertise, as they can increase "the MMN amplitude and/or shorten its latency, both amplitude increment and latency shortening being indices of facilitated formation of neural sound representations in the human brain" (Milovanov et al. 2008). Furthermore, the MMN relies on the "ability to detect the actual physical difference in sensations, and … the extraction of patterns in sound and their relative probabilities … are dependent on the formation and short-term maintenance of a memory trace for the standard and deviant sound" (Fitzgerald and Todd 2020). This ability is a part of the working memory (WM).

The WM is "a brain system responsible for temporary storage and simultaneous manipulation of information, which is critical for higher cognitive functions such as planning, problem solving, and reasoning, but also for understanding or appreciating speech and music" (Schulze and Koelsch 2012, 229). The WM for auditory information has been thought of as a single system encompassing all auditory input, and it was assumed that neural structures engaged with verbal information also engaged with tonal information (Schulze et al. 2011, 771). However, recent functional magnetic resonance imaging studies suggest that there exist two WMs, one for verbal information and one for tonal information. The verbal WM consists of a phonological loop, which is the verbal short-term memory that enables humans to process and rehearse speech stimuli. The tonal WM consists of a tonal loop, which processes and rehearses tonal stimuli. In a study comparing the verbal and tonal WM of non-musicians and musicians it was found that while verbal and tonal WM share core structures, there were "significantly different structural weightings" between the groups (Schulze et al. 2011, 771). Both groups showed an overlap in structures for both verbal and tonal WM, but it was greatest for the musicians, who activated additional areas for tonal WM that were strongly activated in nonmusicians for verbal WM. The musicians were the only group to recruit "a number of structures exclusively for the rehearsal in either of the two domains" (Schulze et al. 2011, 779). However, the authors did not examine whether these additional brain structures influenced the linguistic abilities of musicians, which could be an area for future studies.

In the MMN experiment by Milovanov et al. (2008) the stimuli varied in pitch and consisted of the same stimuli used in the chord discrimination test as mentioned above. During the EEG recordings the sound stimuli were delivered via headphones and the participants were asked to watch a silent movie without subtitles in order to distract their attention from the sound stimuli and to avoid word processing (Milovanov et al. 2008). The results of the MMN study showed no significant difference between the less-advanced pronunciation group (LPG) and the advanced pronunciation group (APG) in the stimulus with the major third being 2% flat, but a significant difference was found between the groups with the major third being 4% flat. Here, the APG showed larger fronto-central responses than the LPG. The authors suggest that this contrast is due to a "difference in the activation of attentional frontal networks", which are more readily activated in participants with high pronunciation skills, and they propose that such activation is also triggered by music (Milovanov et al. 2008).

With the results showing differences in pitch processing of musical sounds between the two groups, the authors wanted to see whether this difference also existed for duration differences in speech and musical sounds in their 2009 study. Duration was chosen because it carries semantic information in vowels and consonants in quantity languages such as Finnish. The duration feature poses a problem when learning a non-quantity language, as the feature is easily transferred into the L2 phonemic system. The authors hypothesized that musical aptitude would be less important than L1 influence (Milovanov et al. 2009). In their study, the participants from the 2008 study were used to examine whether the APG was better than the LPG at processing duration changes in music and speech sounds. In the MMN experiment, the music stimuli consisted of a violin tone with a standard duration of 250 ms and a deviant duration of 150 ms. The speech stimuli consisted of the

Finnish vowel /ø/ with a standard duration of 250 ms and a deviant duration of 150 ms. The study found that the "children with advanced pronunciation and musicality skills displayed enhanced MMNs to duration changes in both speech and musical sounds" compared to the LPG, meaning that the APG encoded the duration differences more effectively (Milovanov et al. 2009). Furthermore, right hemisphere lateralization was found in the APG. Traditionally, musical duration processing has been thought to take place in the left hemisphere where language is processed, while the right hemisphere has been thought to process non-durational aspects of music. However, as right-lateralization was found the results did not follow the predictions of left-hemispheric dominance. The authors propose that a cause of this lateralization could be that the length of the stimuli interfered with the experiment, since "shorter durations (20–50 ms) are typical characteristics of speech and induce stronger responses in the left auditory cortex" (Milovanov et al. 2009). Therefore, it seems that the right auditory cortex, and not the left, was more responsible for musical durations, as the longer duration stimuli were processed there (Milovanov et al. 2009). This indicates that there is an overlap in neural processing of speech and musical sound, and that musical aptitude and linguistic skills are interconnected (Milovanov et al. 2008; 2009).

The question now arises of which acoustic features mediate the link between musical aptitude and linguistic skills. This was addressed by Kempe et al. in 2015.

Musical aptitude and non-native speech-sound processing

Kempe et al. (2015) examined which acoustic features could mediate the link between sensitivity to non-native speech-sound processing and musical ability. The study was based on the idea of shared processing of acoustic features of language and music, whereby increased sensitivity to features related to musical ability benefits language skills (Kempe et al. 2015, 350-351). In this study, American English L1 speakers were tested in their sensitivity to Norwegian vowel and tonal contrasts as well as to the non-linguistic acoustic features pitch, spectral and temporal information. Norwegian was chosen because, as a tonal language, it may be easier to learn for musicians, as sensitivity to pitch "may be critical for the processing of lexical tones or pitch accents … [as the] tonal information requires listeners to process pitch values along with the direction of pitch change" (Kempe et al. 2015, 351). The participants' musical ability was measured separately as musical experience and aptitude, with a questionnaire for the former and the music aptitude test Advanced Measures of Musical Audiation (AMMA) for the latter. The participants were tested in six subtests all of which were conducted as AX discrimination tasks. Two of the subtests tested the participants' sensitivity to Norwegian vowel and tonal contrasts. The vowel contrasts were the /i/-/y/ contrast and its short and long vowel variants, neither of which exist in English. The tonal contrast was based on "lexical tones consisting of rising and falling-rising pitch accents, which distinguish minimal pairs of segmentally identical bi-syllabic words. For example, 'Hammer', spoken with the rising tone, is a Norwegian proper noun while 'hammer', spoken with the falling-rising tone, denotes the tool" (Kempe et al. 2015, 352). The remaining subtests tested the participants' sensitivity to temporal information, pitch, spectral information, and "pitch contours comprising non-speech analogues of the tonal contrast" (Kempe et al. 2015, 354). For temporal information eight sound waves with different amplitude rise times were synthesized, with rise times reaching maximum amplitude at 0, 10, 20, 30, 60, 70, 80, and 90 ms. The sounds were paired up so that their rise times differed by 60 ms, e.g., 10 ms and 70 ms. For the pitch test, eight pure tone sound waves ranging from 100 to 3,000 Hz were synthesized, and each tone was contrasted with a counterpart with a 2% higher frequency. For the spectral sensitivity test complex tones comprising three frequencies, e.g., 100 Hz, 700 Hz, and 2,200 Hz were synthesized. The complex tones were paired with a counterpart with a 2% change in either the middle or the high frequency (Kempe et al. 2015, 353).

The authors conducted a zero-order correlation and a mediation analysis. The zero-order correlation confirmed a link between non-native speech-sound processing and musical aptitude, as "the AMMA tonal score was positively correlated with discrimination of the tonal and vowel contrasts, and the AMMA rhythm score was positively correlated with discrimination of the tonal contrast" (Kempe et al. 2015, 356). No significant correlation was found between musical experience and non-native speech-sound processing. The mediation analyses of tonal contrasts and vowel contrasts are summarized in Figures 1a and 1b, respectively.

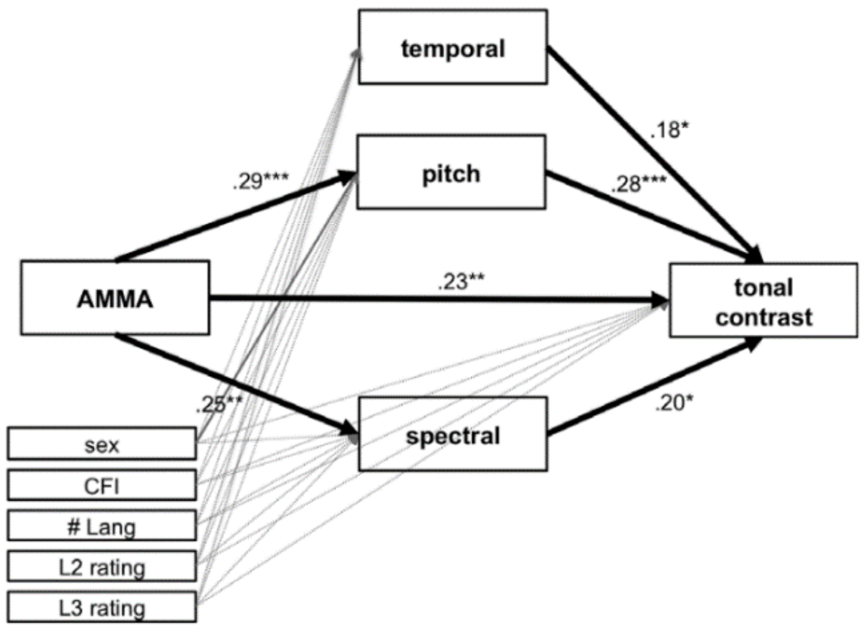

Figure 1a (Kempe et al. 2015, 359). © 2014 The British Psychological Society.

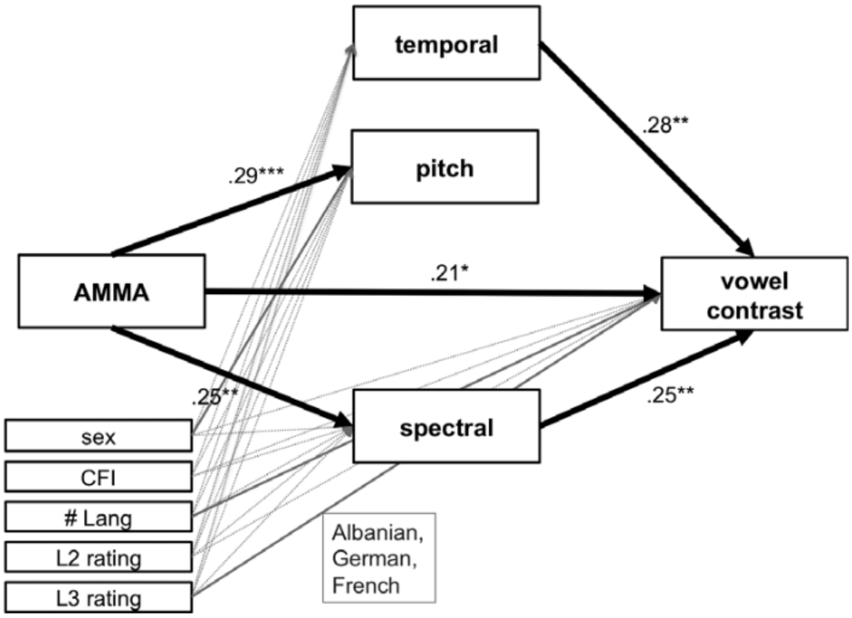

Figure 1b (Kempe et al. 2015, 360). © 2014 The British Psychological Society

The Figures show direct and indirect links between the AMMA score and tonal and vowel contrast processing through temporal, pitch, and spectral sensitivity. Sex, Culture Fair Intelligence (CFI), number of previously learned languages, and self-ratings of L2 and L3 proficiency were controlled for. Only significant links are shown, and the grey lines indicate significant effects of the covariates. Figure 1b also shows which L2s "are most likely responsible for the link between L3-rating and vowel processing" (Kempe et al. 2015, 360). The asterisks indicate significance with ***p < .001; **p < .01; *p < .05.

Both Figures indicate that the AMMA score had a direct effect and an indirect effect, through spectral sensitivity, on the tonal and vowel contrasts. Only in the tonal contrast was an indirect link through pitch sensitivity found. The indirect effects indicate that spectral sensitivity and, for tonal contrasts, pitch sensitivity partially mediated the link between musical aptitude and non-native speech-sound processing (Kempe et al. 2015, 358). Temporal sensitivity was not found to mediate the link between musical aptitude and either contrast. As in the Milovanov et al. 2009 study, this finding was explained by the fact that the time scale for duration change in music is longer than it is for speech, which explains why the temporal stimuli, which were below 100 ms, were processed as speech and not as music (Kempe et al. 2015, 359).

Spectral sensitivity was the most consistent partial mediator in both tonal and vowel contrast analyses. This suggests that "the processing of complex spectral information may be of greater relevance to both music and language than the processing of pure tones" (Kempe et al. 2015, 361). This is unsurprising as multiple frequencies constitute the formants that characterize vowels. Pitch and spectral processing were also found to partially mediate a link for both musical expertise and musical aptitude to non-native speech-sound processing, but only musical aptitude was directly linked and therefore predicted non-native speech-sound processing better than musical experience (Kempe et al. 2015, 359-363).

Conclusion

Musical aptitude was proposed as a better measure of musical ability than musical experience, as aptitude is an inherent sensitivity to features of both musical and speech sounds, and therefore avoids the challenges faced by musical experience in defining "experience" and "musician". However, as many music aptitude tests exist, other alternatives are needed for the accurate measurement of musical ability. The use of neurobiological techniques, e.g., the MMN, might improve the quantification of musical ability, as it measures well-defined brain activity, such as the processing of sound. It was proposed that music and language skills are interconnected as greater overlaps in neural processing of music and speech were found for musicians than nonmusicians, and as individuals with no musical experience but with high L2 production skills were found to have high musical aptitude. Furthermore, spectral sensitivity was found to be the primary partial mediator of the link between musicality and non-native speech-sound processing, with pitch sensitivity having a limited mediating role. The processing of temporal stimuli was found in both Milovanov et al. (2009) and Kempe et al. (2015) to depend on the length of the duration, as stimuli exceeding 100 ms were processed as music and stimuli shorter than 100 ms were processed as speech. It was concluded that sensitivity to temporal features of music did not improve the acquisition of temporal features of L2 vowels. Musical aptitude was found to predict non-native speech-sound processing better than musical experience. As only a partial mediation was found for the link between musical ability and non-native speech-sound processing further studies are needed.

Reference List

Chobert, Julie, and Mireille Besson. 2013. "Musical Expertise and Second Language Learning." *Brain Sci* 3 (2): 923-940. https://doi.org/10.3390/brainsci3020923

Christiner, Markus, and Susanne Reiterer. 2016. "Music, song and speech: A closer look at the interfaces between musicality, singing and individual differences in phonetic language aptitude." In *Cognitive Individual Differences in Second Language Processing and Acquisition*, edited by Gisela Granena, Daniel O. Jackson, and Yucel Yilmaz, 131-156. Amsterdam/Philadelphia: John Benjamins Publishing Company. https://ebookcentral-proquest-com.ez.statsbiblioteket.dk/lib/asb/detail.action?docID=4753485.

Fitzgerald, Kaitlin, and Juanita Todd. 2020. "Making Sense of Mismatch Negativity." *Frontiers in Psychiatry* 11: 468-468. https://doi.org/10.3389/fpsyt.2020.00468.

Gottfried, Terry L. 2007. "Music and Language Learning." In *Language Experience in Second Language Speech Learning: In honor of James Emil Flege*, edited by Ocke-Schwen Bohn and James J. Munro, 221-238. Amsterdam: John Benjamins Publishing Company. https://ebookcentral-proquest-com.ez.statsbiblioteket.dk:12048/lib/asb/reader.action?docID=622748

Kempe, Vera, Dennis Bublitz, and Patricia J. Brooks. 2015. "Musical ability and non-native speech-sound processing are linked through sensitivity to pitch and spectral information." *British Journal of Psychology* 106 (2): 349-366. https://doi.org/10.1111/bjop.12092

Kraus, Nina, and Bharath Chandrasekaran. 2010. "Music training for the development of auditory skills." *Nature Reviews Neuroscience* 11 (8): 599-605. https://doi.org/10.1038/nrn2882

Kraus, Nina, Erika Skoe, Alexandra Parbery-Clark, and Richard Ashley. 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. https://doi-org.ez.statsbiblioteket.dk:12048/10.1111/j.1749-6632.2009.04549.x

Law, Lily N.C., and Marcel Zentner. 2012. "Assessing Musical Abilities Objectively: Construction and Validation of the Profile of Music Perception Skills." *PLoS ONE* 7 (12). https://doi.org/10.1371/journal.pone.0052508

Lidji, Pascale, Pierre Jolicœur, Régine Kolinsky, Patricia Moreau, John F. Connolly, and Isabelle Peretz. 2010. "Early integration of vowel and pitch processing: A mismatch negativity study." *Clinical Neurophysiology* 121 (4): 533-541.

Luck, Steven J. 2014. *An Introduction to the Event-Related Potential Technique.* 2nd ed. Boston: The MIT Press. https://ebookcentral-proquest-com.ez.statsbiblioteket.dk/lib/asb/reader.action?docID=3339822&ppg=1

Milovanov, Riia, Minna Huotilainen, Paulo A.A. Esquef, Paavo Alku, Vesa Välimäki, and Mari Tervaniemi. 2009. "The role of musical aptitude and language skills in preattentive duration processing in school-aged children." *Neuroscience letters* 460 (2): 161-165. https://doi.org/10.1016/j.neulet.2009.05.063

Milovanov, Riia, Minna Huotilainen, Vesa Välimäki, Paulo A.A. Esquef, and Mari Tervaniemi. 2008. "Musical aptitude and second language pronunciation skills in school-aged children: Neural and behavioral evidence." *Brain research* 1194: 81-89. https://doi.org/10.1016/j.brainres.2007.11.042

Milovanov, Riia, Päivi Pietilä, Mari Tervaniemi, and Paulo A.A. Esquef. 2010. "Foreign language pronunciation skills and musical aptitude: A study of Finnish adults with higher education." *Learning and Individual Differences* 20 (1): 56-60. https://doi.org/10.1016/j.lindif.2009.11.003

Nardo, Davide and Susanne Maria Reiterer. 2009. "Musicality and phonetic language aptitude." In *Language Talent and Brain Activity*, edited by Grzegorz Dogil and Susanne Maria Reiterer, 213-255. De Gruyter Mouton. https://doi-org.ez.statsbiblioteket.dk/10.1515/978311021549

Schulze, Katrin, Stefan Zysset, Karsten Mueller, Angela D. Friederici, and Stefan Koelsch. 2011. "Neuroarchitecture of Verbal and Tonal Working Memory in Nonmusicians and Musicians." *Human Brain Mapping* 32 (5): 771-783. https://doi.org/10.1002/hbm.21060

Schulze, Katrin, and Stefan Koelsch. 2010. "Working Memory for Speech and Music." *Annals of the New York Academy of Sciences* 1251 (1): 229-236. https://doi-org.ez.statsbiblioteket.dk/10.1111/j.1749-6632.2012.06447.x

Suzukida, Yui. 2021. "The Contribution of Individual Differences to L2 Pronunciation Learning: Insights from Research and Pedagogical Implications." *RELC Journal* 52 (1): 48-61. https://doi.org/10.1177/003368822098765.

Talamini, Francesca, Massimo Grassi, Enrico Toffalini, Rosa Santoni, and Barbara Carretti. 2018. "Learning a second language: Can music aptitude or music training have a role?" *Learning and Individual Differences* 64: 1-7. https://doi.org/10.1016/j.lindif.2018.04.003.