6-2020

The Impact of Childhood Music Experience on Speech Perception and Processing: A Systematic Review

Erika Lanham

The Graduate Center, City University of New York

THE IMPACT OF CHILDHOOD MUSIC EXPERIENCE ON SPEECH PERCEPTION AND PROCESSING: A SYSTEMATIC REVIEW

by

ERIKA A. LANHAM

A capstone research project submitted to the Graduate Faculty in Audiology in partial fulfillment of the requirements for the degree of Doctor of Audiology, The City University of New York

2020

THE IMPACT OF CHILDHOOD MUSIC EXPERIENCE ON SPEECH PERCEPTION AND PROCESSING: A SYSTEMATIC REVIEW

by

Erika A. Lanham

This manuscript has been read and accepted by the Graduate Faculty in Audiology in satisfaction of the capstone project requirement for the degree of Au.D.

______________________	______________________
Date	Brett A. Martin, Ph.D., CCC-A Faculty Mentor/Advisor
______________________	______________________
Date	Brett A. Martin, Ph.D., CCC-A Executive Officer

THE CITY UNIVERSITY OF NEW YORK

ABSTRACT

THE IMPACT OF CHILDHOOD MUSIC EXPERIENCE ON SPEECH PERCEPTION AND PROCESSING: A SYSTEMATIC REVIEW

by

Erika A. Lanham

Advisor: Brett A. Martin, Ph.D., CCC-A

Objective: The purpose of this investigation was to conduct a systematic review of the literature that addresses the impact of childhood musical experience on speech perception and processing abilities. Specifically, this review assessed how musical training impacted scores on both objective and behavioral tests of speech perception/processing in children. This analysis contributes to a better understanding of the effects of individual musical experience in childhood on our ability to perceive and process speech in a variety of listening conditions. This analysis also determined the clinical implications of such findings.

Methods: A comprehensive search utilizing the Web of Science database accessible through the City University of New York (CUNY) Graduate Center Library was conducted to identify relevant studies published after 2000. Inclusion criteria included the evaluation speech perception and/or processing in children utilizing objective and/or behavioral outcome measures.

Results: Sixteen studies met the inclusion criteria for this systematic review. The studies utilized a variety of outcome measures, which were categorized as objective or behavioral. All included studies found a significant positive relationship between musical experience and speech perception and/or processing abilities in children for both behavioral and objective outcome measures.

Discussion: Significant effects of musical training in childhood were noted across outcome measures suggesting a positive effect on speech perception and processing. Effects on speech perception and processing were noted when both behavioral and objective measures were utilized. Furthermore, studies comparing behavioral and objective outcome measures reported similar findings between the two methods.

Conclusion: The positive effect of childhood musical experience on speech perception and processing abilities is present throughout the literature reviewed when both objective and behavioral outcome measures are utilized. As a result, formal musical training in childhood should be considered as a viable option for auditory training when the goal is improved speech perception and/or processing. The results of these studies should also support the benefit of music classes in school curriculums to help children overcome communication challenges (such as listening in the presence of noise, distance, and poor acoustics) that are frequently found inside and outside of the classroom. Future research should address the limitations of the included studies, such as utilizing a standard musical training program, replicating the large proportion of research on this topic that originated from the Northwestern University Auditory Neuroscience Laboratory, and the utilization of a quasi-experimental or randomized clinical trial design.

Key words: “musical training,” “speech perception/processing,” “listening and learning,” “neuroplasticity,” “auditory processing,” and “children.”

ACKNOWLEDGEMENTS

I would like to extend my deepest gratitude to my esteemed advisor and mentor, Dr. Martin, for her support and guidance throughout my graduate school journey. Your commitment to the growth of your students and the entire Doctor of Audiology program is invaluable.

To my parents and fiancée, I could not have completed this journey without you. Thank you for your unwavering support and love over these past four years. I could never repay you for the sacrifices you have made to allow me to achieve my goals.

TABLE OF CONTENTS

LIST OF TABLES

INTRODUCTION

In recent years, auditory plasticity has gained attention in the audiology community in light of the potential it offers for improved communication abilities. Plasticity is defined as being easily shaped or molded; brain plasticity (or neuroplasticity) is the brain's capability to change itself based on input from the environment via alteration and reorganization of neural pathways and synapses (Hubener & Bonhoeffer, 2014). Auditory plasticity is defined as the auditory system's ability to self-organize the cerebral cortex in response to behaviorally relevant input (Pantev et al., 2006). The concept of auditory plasticity has provided increased understanding of the potential benefits of implementing auditory training activities on speech perception and processing (Moreno & Bidelman, 2013). One relatively new approach is the use of music to potentially improve the processing of sound, including speech.

Musical training is a complex activity, involving somatosensory, motor, visuo-spatial, auditory, executive, and memory functions (Hannon & Trainor, 2007). Studies have shown that when compared to non-musicians, musicians have both structural and functional differences in cortical auditory processing. When compared to non-musicians, musicians showed better temporal acuity for behavioral tasks which can be attributed to rhythm perception, discrimination ability, auditory fusion, representing temporal regularities in performance tasks (Rammsayer & Altenmuller, 2006), and improved ability to reproduce duration intervals (Grondin & Killeen, 2009). Micheyl and colleagues (2006) also illustrated that musicians had significantly lower frequency discrimination thresholds for both complex tones and pure tones when compared to non-musicians, with a larger advantage for the discrimination of harmonic complex tones (in this study, the sum four sinusoids with a fundamental frequency of 300 Hz or 330 + Δf Hz and the corresponding harmonics 2 through 5) than pure tones (in this study, a single sinusoid presented at 330 Hz or 330 + Δf Hz). These findings corroborated previous studies that have demonstrated higher performance in pitch discrimination tasks in musicians via smaller frequency discrimination thresholds (Spiegel and Watson, 1984, Kishon-Rabin et al., 2001). A recent study also suggested that musicians had better frequency change detection abilities than non-musicians in both quiet and noisy conditions; improved frequency change detection ability in musicians was seen in both behavioral tasks and in EEG recordings (Liang et al., 2016).

Musicians have been shown to have structural brain changes; Schlaug and colleagues (1995) reported that musicians had a larger corpus callosum and higher volumes of grey matter in auditory, visuo-spatial, and motor areas. The structural and functional changes outlined indicated that through brain plasticity, the auditory system is processing input more efficiently (Munte, Altenmuller, & Jancke, 2002). Further studies demonstrated structural differences between musicians and non-musicians in areas of the brain involved in both music and communication including areas of the cerebellum, corpus callosum, the anterior-medial portion of Heschl's gyrus, the inferior lateral temporal lobe, the inferior frontal gyrus, the posterior band of the precentral gyrus, and the planum temporale (as cited in Moreno, 2009). These findings pose the question: are these structural and functional changes due to musical training or are individuals with innate brain differences more likely to become musicians? Both Schneider and colleagues (2002) and Hyde and colleagues (2009) reported a strong correlation between the amount of music experience and the magnitude of structural and functional brain changes, suggesting that the above differences could be due to musical training rather than biological predisposition. Schulz and colleagues (2003) also reported increased amplitude of responses and a change in the organization of the primary auditory cortex in event-related potential (ERP) recordings when comparing pre and post musical training measures. Although predisposition for the above outlined structural and functional brain changes cannot be ruled out, there is increased evidence that these brain changes can be attributed to musical experience.

It has been established that both speech and music have within-domain neural plasticity; when you train within one domain, your processing of sound within that one domain is altered. Cross-domain auditory plasticity is the concept that training in one domain (e.g., music) can impact neural processing in another domain (e.g., speech). Patel (2011) proposed a hypothesis for why musical training benefits the neural encoding of speech. The OPERA hypothesis is based on the concept that brain plasticity in speech-processing networks occurs when the following five conditions are met:

(1) **O**verlap: there is overlap in the brain networks that process an acoustic feature used in both speech and music

(2) **P**recision: music places higher demands on these networks than does speech, in terms of the precision of processing

(3) **E**motion: the musical activities that engage this network elicit strong positive emotion

(4) **R**epetition: the musical activities that engage this network are frequently repeated

(5) **A**ttention: the musical activities that engage this network are associated with focused attention (Patel, 2011).

Upon meeting these conditions, the neural networks involved will function at a higher level of precision than typically needed for speech through neural plasticity, thus leading to improved speech processing.

The OPERA hypothesis has since been expanded to not only focus on how music training impacts sensory processing but also cognitive processing and to account for the impact of nonverbal music training. This expansion has been furnished to include research showing enhanced auditory attention and working memory in musicians and overlap in neural networks in the two domains. The expanded OPERA hypothesis proposes that music training enhances speech processing when the following three conditions are met:

(1) A sensory or cognitive process used by both speech and music (e.g., encoding of waveform periodicity; auditory and working memory) is mediated by overlapping brain networks

(2) Music places higher demands on that process than speech

(3) Music engages that process with emotion, repetition, and attention (Patel, 2013).

The basis of the expanded OPERA hypothesis is that higher demands that music places on sensory and cognitive processes shared with speech lead to enhanced speech processing and when combined with emotion, repetition, and attention, lead to neural plasticity (Patel, 2013).

Mechanisms of neural plasticity are believed to arise from changes in neuronal excitability which derive from the interaction of bottom-up inputs and modulation by top-down experience-dependent cortical changes (Tzounopoulos & Kraus, 2009). It has been suggested that corticofugal tuning is the underlying mechanism for the connection between music and language which is supported by efferent pathways from the cortex converging at the midbrain and modifying input. Research by Tzounopoulos & Kraus (2009) supporting this concept suggests that expert listeners who have undergone plastic changes in the auditory system had more efficient corticofugal feedback systems. Additionally, a study by Moreno & Bidelman (2014) showed larger contralateral suppression and less loudness adaptation of Otoacoustic Emissions in musicians, indicating a strengthening of feedback to peripheral auditory processing.

Increased frequency following response magnitude was also reported in musicians for components of speech including fundamental frequency and formants, indicating musicians have finer neural representations of critical components of speech including pitch and timbre (Moreno & Bidelman, 2014). These studies support the concept that musical training strengthens the top-down efferent feedback system. At the level of the cortex, studies have shown enhanced excitability within the primary and secondary auditory cortex and enhanced cortical responses to pitch, timbre, and timing in musicians (Moreno & Besson, 2005). These enhancements in cortical activity are said to manifest as improved responsiveness to speech relevant signals (Moreno & Bidelman, 2014).

A number of auditory training programs have been developed as a method of audiologic rehabilitation based on studies that have shown that neural responses to auditory input can be altered intentionally through intensive listening (Tremblay, Kraus, McGee, Ponton, & Otis, 2001; Tremblay, Shahin, Picton & Ross, 2009; Orduña, Liu, Church, Eddins, & Mercado, 2010). The premise behind auditory training programs is that by exercising the auditory system through sensory input, a person can improve their ability to perceive speech and transfer the skills learned to real-world situations to improve their overall communication ability. Some commonly used computer-assisted auditory training programs include: Computer-Assisted Speech Perception Testing and Training at the Sentence Level (CASPERSent), Computer-Assisted Speech Training (CAST), and Listening & Communication Enhancement (LACE). In general, these programs involve listening to auditory input (typically sentences or phrases) and identifying the words under various conditions such as just auditory input, auditory and visual input, and in noise.

With this understanding of neuroplasticity and the impact that musical training can have on our auditory system's processing ability, it is suggested that musical training can improve speech perception and processing in a similar manner as the aforementioned auditory training programs. The purpose of this study was to conduct a systematic review of the studies on the effectiveness of musical training as a method of improving speech perception and processing abilities in children. Specifically, this review will assess how musical training impacts scores on both objective and behavioral tests of speech perception/processing in children.

METHODS

The Web of Science online database was searched. Search filters included peer-reviewed journals with articles published after 2000. Research published in years prior was not included to ensure a focus on current research given recent renewed interest in this topic in the field. The main search terms utilized were “musical training,” “speech perception/processing,” “listening and learning,” “neuroplasticity,” and “children”. Supplementary search terms, such as, “speech in noise,” “auditory evoked potentials,” and “frequency following response” were utilized to find additional studies for inclusion.

The PRISMA (Preferred Reporting Items for Systematic reviews and Meta-Analyses) statement determined which studies were included in this systematic review. It is designed to improve the quality of studies included in systematic reviews and meta-analyses through the use of a 27-item checklist and four-phase flow diagram (Moher, Liberatie, Tetzlaff & Altman, 2009). This review utilized the following inclusion criteria: articles published in English; typically developing children under the age of 18 that have undergone musical training; and the use of objective outcome measures of speech perception and/or processing such as auditory evoked potentials and/or behavioral outcome measures of speech perception and/or processing such as speech in noise testing, questionnaires, or standardized tests.

In total, the Web of Science database search yielded a total of 195 studies which were then screened to ensure the inclusion criteria were met. 165 studies were excluded on the basis of not meeting inclusion criteria, yielding 30 studies for in-depth assessment. Subsequently applied exclusion criteria were studies with small sample sizes (less than 10 participants), those with poor reporting of protocol or data analysis (such as insufficient detail of methodology used to carry out the research design and insufficient detail of data analysis), and studies with inadequate or poorly described matching of participants between groups, leaving 16 studies to be evaluated in this systematic review.

Articles included in this systematic review were assessed for type and duration of musical training, independent variables, dependent variables, sample size, and results. Studies were also divided into subsections dependent on the type of outcome measures used, either objective or behavioral. Results not relevant to this systematic review, such as data collected on adult subjects and data regarding factors not pertaining to speech perception and/or processing, were not included in this analysis.

RESULTS

The 16 studies evaluated employed a (1) quasi-experimental research design, (2) a descriptive, prospective cohort research design, or (3) a descriptive, retrospective cohort research design. Of the 16 studies, 50% employed a prospective cohort design whereby outcome measures were evaluated before and after musical training (Dittinger et al., 2017; Putkinen et al., 2013; Putkinen et al., 2014; Strait et al., 2012; Strait et al., 2013; Strait et al., 2014; Vasuki et al., 2017; Zuk et al., 2014). 18.75% of studies employed a retrospective cohort design whereby outcome measures were evaluated after or during musical training (Habibi et al., 2016; Slater et al., 2015; Tierney et al., 2013). The remainder 31.25% of studies employed a quasi-experimental design whereby subjects were not randomly selected, however, they were randomly assigned into groups dictating their musical training (Chobert et al., 2012; Gerry et al., 2012; Kraus et al., 2014a; Kraus et al., 2014b; Moreno et al., 2009).

Of the 16 studies, 68.75% were based on a sample size of greater than 30, while the remainder 31.25% of studies were based on a sample size of less than 30, with the smallest sample size included in this evaluation being 19 (Kraus et al., 2014a). All included studies drew from local participants, with 56.25% of studies taking place in the United States, and the remainder 43.75% of studies taking place in France (Chobert et al., 2012; Dittinger et al., 2017), Canada (Gerry et al., 2012), Portugal (Moreno et al., 2009), Finland (Putkinen et al., 2013, Putkinen et al., 2014), and Australia (Vasuki et al., 2017). A majority of the studies (75%) drew participants ages 6-13; Gerry et al., (2012) drew participants who were 6 months of age at baseline and 11.5 months of age at the final data collection; two studies (12.5%) drew participants ages 3-5 years of age (Strait et al., 2013; Strait et al., 2014). Inclusion and exclusion criteria varied between studies; see Table 1 for study design characteristics.

A majority of the studies (68.75%) compared only two groups: children with musical training and children without musical training; 12.5% of studies compared three groups: two groups with different types of musical training and one group with no musical training; 12.5% of studies compared two groups differing in length of musical training; while the remaining study (6.25%) compared a group receiving musical appreciation training and a group receiving training on both an instrument and musical appreciation. See Tables 1 and 2 for further information regarding study design.

Outcome measures to be analyzed include both behavioral and objective tests of speech perception and processing. Of the 16 studies, 2 (12.5%) implemented only behavioral outcome measures; 7 (43.75%) studies implemented only objective outcome measures; and 7 (43.75%) implemented both behavioral and objective outcome measures. Behavioral outcome measures included scores on various tasks such as phonological categorization, tonal/rhythm discrimination, pitch discrimination, speech-in-noise, auditory processing, and cognitive assessments. Objective outcome measures included analysis of auditory evoked potentials such as Mismatch Negativity (MMN), speech evoked ABR (cABR), P1-N1-P2 complex, and P3a. Of the 14 (87.5%) of studies who implemented objective outcome measures, 18.75% utilized MMN, 36% utilized cABR, 18.75% utilized P3 or P3a, 6.25% utilized fMRI and 25% utilized other various EEG recordings. See Table 2 for types of outcome measures used in each study and Table 5 for a summary of the significant findings in each study, divided by type of outcome measure.

While all the studies included in this analysis controlled for extraneous variance to some degree, only 50% of the studies implemented a standardized musical training program over a determined time period. The remaining 50% of studies employed questionnaires to limit the variance between length and rigor of the musical training. For those studies who did not implement a standardized musical training program, their method of quantifying the participants' musical training can be seen in Table 3.

Description of Auditory Evoked Potentials Utilized

Auditory Evoked Potentials (AEPs) can be divided into multiple categories. Picton (2010) suggests categorizing AEPs based on the latency of the components: early, middle, and late. The early AEPs occur between 1-15 ms after stimulus onset and are comprised of the Auditory Brainstem Response (ABR) waves I-VIII, Frequency Following Response (FFR), and the Auditory Steady-State Response (ASSR). The middle AEPs occur between 10-50 ms after stimulus onset and contain the Middle Latency Response (MLR) which is comprised of peaks Na, Pa, and Nb. The late AEPs occur between 50-500 ms and contain the P1-N1-P2 complex, Mismatch Negativity (MMN), and P3 (or P300) components. An alternative method of categorization, as seen in Steinschneider and Dunn (2002), is based on how the AEPs are processed/obtained: sensory-evoked/exogenous or processing-contingent/endogenous. The sensory-evoked AEPs do not require subject attention as the responses are evoked by the various physical attributes of the stimulus; the sensory-evoked AEPs include the ABR, FFR, ASSR, and P1-N1-P2 complex. The processing-contingent AEPs require further perceptual and/or cognitive processing (either via automatic functions not requiring subject attention or via attention-dependent and active processing); the processing-contingent AEPs include MMN and P3.

In order to interpret the findings of the AEP research included in this systematic review, one must first understand the methods used to obtain the measurements, the waveform composition, where the response is generated within the auditory system, and the implications of such findings. See Table 4 for a description of the AEPs utilized in this systematic review (Martin et al., 2008; Picton, 2010; Sussman et al., 2013).

Table 1: Study Design Characteristics

Study	Size (N)	Participants	Inclusion/Exclusion Criteria	Statistical Analysis
(Chobert et al., 2012)	24	French children aged 8-10 years; musician group (n=12), painting group (n=12)	Native French speakers, no known deficits, similar SES[a], no pre-test musical or painting training	Five-way repeated-measures ANOVA[b], four-way ANOVA, Greenhouse-Geisser corrections applied when appropriate, Tukey post-hoc test
(Dittinger et al., 2017)	23	French children aged 8-12 years; musician group (n=12), non-musician group (n=11)	Native French speakers, no known hearing or neurological deficits	Two-way ANOVA, 2x3x3x3 ANOVA, 2x2x3x3 ANOVA, Tukey post-hoc test
(Gerry et al., 2012)	60	Canadian infants, average age of 11.5 months at time of final testing; active music group (n=20), passive music group (n=14), no music group (n=26)	Similar SES, no pre-test musical training	ANOVA
(Habibi et al., 2016)	37	Los Angeles-based children aged 6-7 years; music group (n=13), soccer group (n=11), no training group (n=13)	Similar SES, raised in bilingual households, fluent English speakers, no known developmental or neurological disorders	ANOVA, repeated measures ANOVA, 3x3 ANOVA, Greenhouse-Geisser corrections, Tukey post-hoc test
(Kraus et al., 2014a)	19	Los Angeles-based children aged 7-10 years; music appreciation group (n=10), music appreciation + instrument playing group (n=9)	Normal audiological screening, scores within normal limits on perceptual, cognitive, and neurophysiological tests	MANOVA[c], Chi-Square
(Kraus et al., 2014b)	44	Los Angeles-based children aged 6-9 years; 1 year of musical training group (n=18), 2 years of musical training group (n=26)	Similar SES, normal latency response to click-evoked ABR[d] at baseline	Arithmetic mean, repeated measures ANCOVA[e], Bonferroni post-hoc test
(Moreno et al., 2009)	32	Portuguese 8 year olds; musical training group (n=16), painting group (n=16)	No pre-test formal musical or painting training, normal hearing, right-handed, native speakers of Portuguese, similar SES	ANOVA

(Putkinen et al., 2013)	133	Finnish 7-13 year olds; instrumental training group, no musical training group (number of participants varies based on time of data collection)	Similar SES, no known hearing or neurological impairments	3x2 repeated measures ANOVA, independent sample t-test
(Putkinen et al., 2014)	117	Finnish 7-13 year olds; instrumental training group, no musical training group (number of participants varies based on time of data collection)	Similar SES, no known hearing or neurological impairments	T-tests, Bonferroni corrected pairwise post hoc test
(Slater et al., 2015)	38	Los Angeles- based 8 year olds, 1 year of musical training group (n=19), 2 years of musical training group (n=19)	Similar SES, hearing within normal limits, no known learning or neurological impairments, no prior musical training	Repeated-measures ANCOVA post hoc paired t-test, one-way ANOVA, Pearson correlation, Bonferroni correction
(Strait et al., 2012)	31	7-13 year olds, instrumental musical training group (n=15), no musical training group (16)	Hearing within normal limits, normal wave V click-evoked ABR latencies	One-way ANOVA, repeated measures ANOVA, independent sample t-test
(Strait et al., 2013)	26	3-5 year olds, musical training group (n=13), no musical training group (n=13)	Hearing within normal limits, no known neurological or developmental abnormalities, normal wave V click-evoked ABR latencies, normal verbal IQ (measured by PPVT[f])	Repeated measures ANOVA, independent sample t-test, Mann-Whitney U test
(Strait et al., 2014)	47	3-5 year olds, musical training group (n=12), no musical training group (n=9); 7-13 year olds, musical training group (n=13), no musical training group (n=13)	Hearing within normal limits, no known neurological or learning deficits, normal wave V click-evoked ABR latencies and normal IQ (measured by PPVT in 7-13 year olds, WASI[g] in 3-5 year olds)	Repeated-measures ANOVA, post hoc-independent sample t-test, 1-way ANOVA
(Tierney et al., 2013)	43	Chicago-based high-school students (age 14 at pre-test), musical training group (n=21), fitness training group (n=22)	Little to no formal musical training, similar SES, hearing within normal limits, normal click-evoked ABR latencies, no diagnosis of a reading	Repeated measures ANOVA, one-tailed post hoc paired t-tests, one-tailed t-tests

			disorder	
(Vasuki et al., 2017)	50	9-11 year olds, musical training group (n=25), no musical training group (n=25)	Native English speakers, normal or corrected-to-normal vision, hearing within normal limits, present OAEs[h], no known language/reading/cognitive impairments, similar SES	MANOVA, one-sample t-test, independent sample t-test, ANOVAs, MANCOVA[i], arithmetic mean
(Zuk et al., 2014)	27	9-12 year olds, musical training group (n=15), no musical training group (n=12)	Similar SES, similar IQ (measured by D-KEFS[j]), no known neurological or psychological disorders, no head injuries, normal vision and hearing	Independent t-test, independent two-sample t-test

[a] Socioeconomic status
[b] Analysis of Variance
[c] Multivariate Analysis of Variance
[d] Auditory brainstem response
[e] Analysis of Covariance
[f] Peabody Picture Vocabulary Test
[g] Wechsler Abbreviated Scale of Intelligence
[h] Otoacoustic emissions
[i] Multivariate analysis of covariance
[j] Delis–Kaplan Executive Function System

Table 2: Study Variables and Procedures

Study	Independent Variable(s)	Outcome Measure(s)	Procedure
(Chobert et al., 2012)	1. Groups: music, painting 2. Session: T0 (baseline), T1 (after 6 months), T2 (after 12 months)	**MMN[a] amplitude for frequency, duration, & VOT[b] deviants (each with two levels of deviance: small and large); measured at T0, T1, & T2.	Participants watched a silent subtitled movie during EEG[c] recordings from 32 active Ag-Cl electrodes placed according to the 10/20 System using a Biosemi amplifier system. Stimuli presented through headphones.
(Dittinger et al., 2017)	Groups: MUS (musical training), NM (no musical training)	*Phonological categorization task, word learning (phase 1 and 2), matching task, semantic task. **EEG data measuring N100 amplitude, N200 and N400 mean amplitudes throughout behavioral tasks.	Participants completed the tasks during which EEG was recorded from 32 active Ag-Cl electrodes placed according to the 10/20 System using a Biosemi amplifier system. Stimuli presented through headphones.
(Gerry et al., 2012)	1. Groups: Active Training, Passive Training, and No Training 2. Time: pre-training, post-training	*Measure of sensitivity to western tonality, measure of social-emotional development, measure of early communicative development.	Participants completed the tonality task (post training) on his/her parent's lap in a sound attenuating chamber, with stimuli presented via two speakers; social-emotional development was examined (pre and post training) via parental report using the IBQ[d]; early communicative development was examined (pre and post training) via the MB-CDIs[e] .
(Habibi et al., 2016)	1. Groups: music group, soccer training, no training 2. Time: baseline, year 2	*Active tonal/rhythm discrimination task (same/different judgement by participant). **Passive tonal perception task (EEG measuring P1, N1, P2); EEG measuring P2, N2, P3 during active task.	Auditory stimuli delivered binaurally via ER-3 insert earphones; EEG continuously recorded. Passive task was completed while watching a silent movie. Active task required participants to determine if stimuli are same or different with a button press

			response.
(Kraus et al., 2014a)	Group: Mus (1 year of music appreciation), Mus+Inst (.5 year of music appreciation, .5 year of instrumental classes)	**cABR[f] to square-wave click and /d/ measuring latency and spectral components.	cABR presented via an ER-3A insert earphone in the right ear using an Intelligent Hearing Systems SmartEP system.
(Kraus et al., 2014b)	1. Group: group 1 (1 year of music training), group 2 (2 years of music training) 2. Year: baseline, year 1, year 2	**cABR to /ba/ and /ga/, analyzed using a cross-phaseogram procedure (quantifying the difference in distinction between /ba/ and /ga/).	cABR presented via an insert earphone in the right ear using an Intelligent Hearing Systems SmartEP system.
(Moreno et al., 2009)	1. Group: music, painting 2. Time: pre-training, post-training	*Musical and speech pitch discrimination tasks. **Continuous EEG recordings throughout tasks analyzed for mean amplitude and latency.	EEG recordings from 32 active Ag-Cl electrodes placed according to the 10/20 System using a Biosemi amplifier system. For pitch discrimination tasks, participants determined if the last word/note of the stimuli was normal or strange via button press.
(Putkinen et al., 2013)	1. Group: music, control 2. Age: 7, 9, 11, 13	**EEG recordings (measuring MMN and P3a) during two oddball paradigms: a chord paradigm, and multi-feature paradigm.	EEG recordings from either (1) a Neuroscan system using 9 Ag-AgCl electrodes, or (2) a BioSemi Active-Two system using 64 Ag-AgCl placed according to the international 10/20 system.
(Putkinen et al., 2014)	1. Group: music, control 2. Age: 9, 11, 13	**EEG recordings (measuring MMN and P3a) during a melodic multi-feature paradigm.	EEG recordings from either (1) a Neuroscan system using 9 Ag-AgCl electrodes, or (2) a BioSemi Active-Two system using 64 Ag-AgCl placed according to the 10/20 system.
(Slater et al., 2015)	1. Group: group 1 (1 year of musical training), group 2 (2 years of musical training) 2. Year: baseline, year 1, year 2	*HINT[g]	HINT administered in a quiet room via Sennheiser HD 25-1 headphones using the standard HINT protocol for children.

(Strait et al., 2012)	Groups: Mus (musicians), NonMus (non-musicians)	*HINT, WIN[h], IVA[i], WJ-III-COG[j] AWM[k] subtest, Colorado Assessment Tests 1.2 - VWM[l] subtest. **cABR /da/ in quiet and in the presence of multi-talker babble with a +10 dB SNR[m], measuring timing, stimulus-to-response fidelity, and spectral encoding.	HINT and WIN administered in a soundproof booth. IVA Test administered in a soundproof booth via Sennheiser HD 25-1 headphones. AWM and VWM tests conducted on a computer. cABR presented via ER-3 insert earphones using NeuroScan Acquire 4.3 equipment and Ag-AgCl electrodes.
(Strait et al., 2013)	1. Group: Mus (musicians), NonMus (non-musicians) 2. Time: year 1, year 2	**cABR /da/ in quiet and in the presence of multi-talker babble at a +10 dB SNR, measuring latency and amplitude.	cABR presented via an ER-3 insert earphone in the right ear using NeuroScan Acquire 4.3 equipment and Ag-AgCl electrodes.
(Strait et al., 2014)	1. Group: Mus (musicians), NonMus (non-musicians) 2. Age: preschoolers (3-5 years old), school-aged children (7-13 years old)	*3-5 year olds: IQ (PPVT[n]); 7-13 year olds: IQ (WASI[o] verbal and nonverbal subtests), AWM (AWM subtest of the WJ-III-COG), VWM (Colorado Assessment Tests 1.2 - visual span subtest), auditory and visual attention (IMAP test - attention subtests). **All participants: cABR to /ga/ and /ba/ stimuli pseudorandomly within the context of 6 other syllables, measuring phase shifts.	IQ, working memory, and attention tasks administered using standard protocol; cABR presented via ER-3 insert earphones using NeuroScan Acquire 4.3 equipment and Ag-AgCl electrodes.
(Tierney et al., 2013)	1. Group: music training, fitness training 2. Time: pre-training, post-training	**cABR /da/ in the presence of multi-talker babble at a -10 dB SNR, employing stimulus-to-response correlation and cross-phaseogram analysis measuring neural response timing.	cABR presented via an ER-3 insert earphone in the right ear using the NeuroScan Stim2 equipment and Ag/Ag-Cl electrodes. Recorded in a sound-attenuated chamber.
(Vasuki et al.,	Group: musicians, non-musicians	*Auditory processing tasks (Musical Ear test - melody and rhythm subtests,	Behavioral auditory and cognitive tasks administered in a sound-treated

2017)		frequency discrimination task, Dichotic Digits test), statistical learning tasks (auditory and visual embedded triplet tasks) with a test phase during which participants indicated which of the two triplets were familiar via a button press. **During the familiarization phase of embedded triplet tasks: EEG was recorded (measuring P100, N250, N200, and P300).	booth on session day 1. On session day 2, EEG was recorded using the Neuroscan system, with 64 Ag/AgCl electrodes placed according to the international 10-20 system.
(Zuk et al., 2014)	Group: musically trained, untrained	*Cognitive assessment (D-KEFS[p] - Trail Making, Verbal Fluency, and Color-Word Interference subtests, WISC IV[q] - Coding subtest, KBIT[r]), set-shifting task. ** fMRI[s] during set-shifting task.	Cognitive assessment tasks administered using standard procedure; set -shifting task required responses via a button press; two fMRI runs conducted on a Siemens 3 T Trio scanner.

[a] Mismatch negativity
[b] Voice onset time
[c] Electroencephalogram
[d] Infant Behavior Questionnaire
[e] MacArthur-Bates Communicative Development Inventories
[f] Speech-Evoked Auditory Brainstem Response
[g] Hearing in Noise Test
[h] Words in Noise Test
[i] Integrated Visual and Auditory Test of Auditory working memory
[j] Woodcock Johnson III Test of Cognitive Abilities
[k] Auditory working memory
[l] Visual working memory
[m] Signal-to-noise Ratio
[n] Peabody Picture Vocabulary Test
[o] Wechsler Abbreviated Scale of Intelligence
[p] Delis-Kaplan Executive Function System
[q] Wechsler Intelligence Scale for Children, Fourth Edition
[r] Kaufman Brief Intelligence Test
[s] Functional Magnetic Resonance Imaging
* Behavioral outcome measure
** Objective outcome measure

Table 3: Quantification of Musical Training in Studies without Standardized Training

Study	Method of Quantification	Musical Training
(Dittinger et al., 2017)	Parental report	Participants practiced music for an average of 4.9 years (range 4-7 years); piano (5), trumpet (2), trombone (2), violin (2), saxophone (1)
(Putkinen et al., 2013)	Parental questionnaire	Participants started playing an instrument approximately at age 7 and receive training in school; most common instruments: violin, viola, cello, double bass, guitar, flute
(Putkinen et al., 2014)	Parental questionnaire	Participants started playing an instrument approximately at age 7 and receive training in school; most common instruments: violin, viola, cello, double bass, guitar, flute
(Strait et al., 2012)	Parental questionnaire	Participants started playing an instrument by age 5, are currently undergoing private training, and have consistently practiced for at least 4 years
(Strait et al., 2013)	Parental report	Participants are currently undergoing private or group musical training, and have been doing so for at least 12 consecutive months; types of training include Kindermusik, Music Together, and Orff music classes
(Strait et al., 2014)	Parental report (3-5 year olds), self report (7-13 year olds)	3-5 year olds: currently undergoing consistent musical training for a minimum of 12 months, group or private training, varying methods 7-13 year olds: began training by age 6 with consistent practice for at least 3 years, varying methods
(Vasuki et al., 2017)	Parental questionnaire	Average of 3.9 years of private musical training with varying methods
(Zuk et al., 2014)	Parental report	At least 2 years of private instrumental training, beginning on average at age 5; piano (5), strings (5), woodwinds (2), guitar (1), percussion (2)

Table 4: Description of Auditory Evoked Potentials Utilized

AEP	Measurement	Composition	Generators	Implications
cABR	No behavioral task or subject attention is required. Auditory stimuli, most commonly used is the syllable /da/, are repeatedly presented. ERPs time locked to the stimulus are measured via scalp electrodes.	Onset waves V (same as wave V of the tonal ABR) and A, followed by consonant-vowel transition wave C, followed by Frequency Following Response waves D through F, followed by offset wave O	Same as the tonal ABR. Multiple brainstem generators including the cochlear nucleus, superior olivary complex, lateral lemniscus, and inferior colliculus.	Reflects the encoding of the fundamental frequency and harmonic structures of speech stimuli. An objective measure of language encoding ability.
P1-N1-P2 Complex	No behavioral task or subject attention is required. Auditory stimuli are repeatedly presented. ERPs time locked to the stimulus are measured via scalp electrodes.	A peak (P1), followed by a trough (N1), followed by a peak (P2).	P1: Heschl's gyrus, hippocampus, lateral temporal regions, and possibly subcortical regions. N1: Multiple generators within primary and secondary auditory cortex including the superior temporal lobe and superior temporal gyrus. P2: Multiple generators including the primary and secondary auditory cortex and the mesencephalic reticular activating system.	Reflects processing of the spectro-temporal feature changes within auditory stimuli. An objective measure of auditory encoding (detection).
MMN	No behavioral task or subject attention is required, although attention can be added to modify the MMN response. A series of auditory stimuli are	Similar composition of the P1-N1-P2 complex with either the addition of a trough (N2) after P2, an enhancement of N1, or an attenuation of P2.	Primary and secondary auditory cortex (believed to index change detection), and possibly frontal cortex (believed to index attention-switching).	Reflects auditory change detection; can index the brain's ability to distinguish context-based changes in a standard repeating regularity.

	presented in a random sequence with “standard” or frequent stimuli and “deviant” or infrequent stimuli. ERPs time locked to the deviant stimulus are measured via scalp electrodes.			A cortical representation of auditory scene analysis as well as an objective measure of central auditory processing.
P3/P300 (comprised of P3a and P3b waves)	Subject attention is required. A series of auditory stimuli are presented in a random sequence with “standard” or frequent stimuli and “deviant” or infrequent stimuli. ERPs time locked to the deviant stimulus are measured via scalp electrodes.	A large peak (P3) after P2.	Widespread generators including auditory cortex, centroparietal cortex, hippocampus, and frontal cortex. P3a: Frontal scalp distribution. P3b: Centroparietal scalp distribution.	Believed to measure the length of time spent processing stimuli and/or the further processing of consciously discriminated sounds.

Table 5: Significant Findings, Divided According to Type of Outcome Measure(s) Utilized

Study	Objective Outcome Measures	Behavioral Outcome Measures
(Chobert et al., 2012)	Group-by-session interaction for duration deviants; music group MMNs were larger at T2 (-1.54 µV) than T1 (-0.30 µV; P<0.05). Main effect of session (P<0.02) and group-by-session interaction (P<0.05) for VOT deviants; music group MMNs were larger at T2 (-3.98 µV) than T1 (-2.14 µV; P<0.02), and marginally larger at T2 than T1 (-2.37µV; P=0.07).	N/A
(Dittinger et al., 2017)	Modulation of N2 and N400 amplitude during word learning phase 1 (Block 1 vs 2, p<0.05; Block 2 vs 3, p<0.01) in musicians. Larger amplitude of N400 (avg. -1.40 µV, p=0.05) and N2 (avg. -2.16 µV, p=0.05) during matching task in musicians. Larger N400 amplitude (avg. -1.37 µV, p=0.04) and N2 amplitude (avg. -1.40 µV, p=0.04).	Phonological categorization task: Main effect of group (p=0.01); musicians made fewer errors (9.7%) than non-musicians (17.2%). Matching task: Main effect of group (p=0.03); musicians made fewer errors (18.2%) than non-musicians (28.9%). Semantic task: Main effect of group (p=0.05); musicians made fewer errors (23.1%) than non-musicians (33.9%).
(Gerry et al., 2012)	N/A	Western tonality task: Difference between groups on proportion of time looking (p=0.05); looking proportions were different from chance (proportion of .5) in the active training group (M=.55, p=0.02) only. Social-emotional development: Infants in active music class showed less distress to limitations (p=0.001), less distress to novel stimuli (p<0.001), more smiling and laughter (p<0.001), and were more easily soother (p<0.001) than those in the passive music class. Early communicative development: Greater increase in use of gestures (p=0.01) in active

		music group (mean score difference of ~20 pre vs post training) than passive music group (mean score difference of ~12 pre vs post training).
(Habibi et al., 2016)	Tonal perception task: Group-by-year interaction; decrease in P1 amplitude from baseline to year 2 in music group (p=0.02). Larger relative percent amplitude difference between P1 and N1 in music group; M +/- SD = -23+/-43.6% for music group (p=0.03). Group difference in P1 amplitude at year 2 (p=0.02); post-hoc contrast was only sig. between music (avg. latency 84 ms) and sports (avg. latency 92 ms) groups (p=0.01). Pitch discrimination task: Difference in P3 amplitude between groups (p=0.01); larger P3 in music than no-training (p=0.01) and sports (p=0.10).	Music group showed higher accuracy in detecting pitch changes (p=0.001); mean score of 62% accuracy in music group, 30.4% in sports group, 45.1% in no training group.
(Kraus et al., 2014a)	Faster response timing in music plus instrument group (p<0.05) for peaks V (avg. 0.46 ms faster), E (avg. 0.81 ms faster), and F (avg. 0.74 ms faster). Trending group effect (p<0.1) for peak A (avg. 0.59 ms faster) and a stronger representation of high harmonics in music plus instrument group.	N/A
(Kraus et al., 2014b)	Group-by-year interaction (p=0.029); improved distinction of contrastive speech sounds in the 2 years of musical training group. Correlation between hours of music training and change in neurophysiological distinction (r=0.481, p=0.001).	N/A
(Moreno et al., 2009)	Session by congruity interaction (P<0.001) for music task; enhanced N300 amplitude to weak incongruities after musical training at midline (avg. 6.27 µV) and lateral (avg. 5.07 µV) electrodes in musicians. Session by congruity interaction (P<0.001) for speech task; enhanced 200-900 ms positive component amplitude to	Reading task: Improvement in percentage of errors after training for music group in the inconsistent condition (group by session by word type interaction, P<0.005); ~30% improvement in music group, ~10% improvement in painting group.

	weak incongruities after musical training at midline (avg. 5.29 µV) and lateral (avg. 4.72 µV) electrodes in musicians.	Pitch (music) discrimination task: Trending improvement post training for weak incongruities in music group (mean improvement 15%, $P<0.006$) but not painting group (mean improvement 5%, $P>0.20$). Pitch (speech) discrimination task: Group by session by congruity interaction ($P<0.02$); improvement post training for weak incongruities in music group (mean improvement 19%, $P<0.001$) but not painting group (mean improvement 9%, $P>0.70$).
(Putkinen et al., 2013)	Main effect of group ($p<0.001$); music group MMN amplitude was larger across all age groups. Group-by-age interaction for chord MMN and chord P3a ($p<0.05$); music group MMN and P3a amplitude increased more steeply with age.	N/A
(Putkinen et al., 2014)	Age-by-group interaction ($p<0.05$); music group MMN increased more steeply with age in response to melody, rhythm, timbre, and tuning modulations.	N/A
(Slater et al., 2015)	N/A	Group by year interaction ($p=0.022$). Greater mean SNR change in 2 years of training group (-2.1dB, $p=0.001$) than in 1 year of training group (no SNR change). Relationship between total hours of training and HINT performance, with more hours linked to better HINT performance ($r=-0.448$, $p=0.005$).
(Strait et al., 2012)	Group-by-condition interaction ($p<0.02$), with the musician group showing less response degradation in noise than non-musicians ($p=0.005$). Earlier latencies of formant transition peaks in quiet and noise for musician group ($p<0.05$). Main effect of group; musicians had more robust representation of harmonics ($p<0.01$).	HINT: Better performance on spatially separated HINT ($p<0.01$) for musicians (mean scores 45%) than non-musicians (mean scores 30%). Working memory: Better performance on auditory working memory task ($p<0.05$) for musicians (mean scores 122) than non-musicians (mean scores 110).
(Strait et al.,	Main effect of group ($p=0.004$); earlier latencies of	N/A

2013)	onset and transition peaks in both quiet and noise conditions for musicians (p<0.05). Less timing delays in noise for musicians (p=0.02).	
(Strait et al., 2014)	Main effect of time range (P=0.03); musicians showed greater positive phase shifts for formant transitions (p<0.01).	Working memory: School-aged musicians outperformed non-musicians on auditory working memory task (P=0.01); mean score of 128.4 for musicians, 116.6 for non-musicians. Attention: School-aged musicians outperformed non-musicians on auditory attention task (P=0.05); mean score of 391.0 for musicians, 489.9 for non-musicians.
(Tierney et al., 2013)	Interaction between year and musical training; decreased stimulus-response lag for musicians post training (avg. shift -0.25 ms, p=0.028) compared to controls (avg. shift 0.14, p=0.239).	N/A
(Vasuki et al., 2017)	Larger triplet onset effect during aSL (p<0.05) and vSL (p<0.005) task in musicians.	Auditory processing: Musicians outperformed non-musicians on the following tasks (p<0.01): melody discrimination (mean difference 6.8%), rhythm discrimination (mean difference 5.8%), music score (mean difference 6.3%), and frequency discrimination test (log) (mean difference -0.6). Statistical learning: Musicians outperformed non-musicians in the auditory statistical learning task (p<0.005); musician mean score of 68.9%, non-musician mean score of 54.7%. Moderate positive correlation between auditory statistical learning scores and music scores (r=0.41, p<0.005).
(Zuk et al., 2014)	Greater activation in the left ventrolateral prefrontal cortex (p=0.005) and bilateral supplementary motor area (p=0.048) in the music group.	Better performance in musically trained children on Coding (mean difference 1.96, p=0.012), Verbal Fluency (mean difference 2.63, p=0.016), and Trail Making (mean difference 2.00, p=0.026).

Objective Outcome Measures

Chobert et al., (2012) found statistically significant enhancements of MMN amplitude in response to duration and VOT deviants after 12 months of musical training while no significant change in MMN amplitude was exhibited after 12 months of painting training (See Table 4 for significant findings). It was also indicated that after only 6 months of musical training, no significant change in MMN amplitude was seen for the duration deviants and a marginal change was seen for the VOT deviants. These findings were said to indicate that at least 6 months of musical training is required to improve pre-attentive processing. Similarly, Putkinen et al., 2013 and Putkinen et al., 2014 found significant enhancements of MMN amplitude in response to a chord paradigm and melodic multi-feature paradigm, respectively, in musically trained children and not their untrained counterparts. Enhanced MMN amplitude was considered to be indicative of heightened sensitivity to changes in auditory stimuli, and represents pre-attentive processing of auditory input.

Kraus and colleagues published two studies measuring the impact of musical training on speech perception/processing via cABR (Kraus et al., 2014a; Kraus et al., 2014b). The first study compared the effect of 1 year of musical appreciation versus half a year of musical training followed by half a year of learning to play an instrument; results indicate stronger neural processing of speech in the musical training plus instrumental learning group due to decreases in latency and more robust spectral representation of components of the cABR waveform (Kraus et al., 2014a). The second study compared the effect of 1 year of musical training to 2 years of musical training; results indicated improved distinction of similar speech sounds in the 2 years of training group illustrated by the stronger responses on cross-phaseogram difference plots (Kraus et al., 2014b). These results were said to demonstrate improved neurophysiological distinction of contrastive speech syllables as a result of musical training and that more training leads to greater improvement, suggesting that musical training influenced auditory processing.

In a study comparing cABR waveforms of musicians and non-musicians presented both in quiet and in noise conditions, musicians demonstrated: less response degradation in noise, earlier response latencies in quiet and noise, and more robust representation of harmonics in quiet and noise compared to their non-musician counterparts (Strait et al., 2012). A similar study that implemented cABR also found earlier latencies to speech onsets and formant transitions in quiet and in noise, decreased timing shifts in noise, and decreased onset peak degradation in noise in musicians compared to non-musicians (Strait et al., 2013). A final study by the same researcher analyzed subcortical encoding of speech signals in musicians compared to non-musicians via cABR and found that musicians demonstrated greater positive phase shifts for formant transitions, which was considered to indicate more temporally distinct neural responses to similar speech signals (Strait et al., 2014). Tierney et al., (2013) analyzed the impact of high school music classes on speech processing via cABR; findings indicated enhanced neural representation of speech in the presence of noise in those enrolled in music class, shown by decreased lag in neural response timing after 1 year of musical training when a cross-correlational analysis was used.

The nomenclature utilized for the late AEP components varied between researchers. Moreno et al., (2009) labeled the AEP waveforms in accordance with their latency, naming the waveform peak with a latency around 250 ms N250 (more commonly referred to as N2). Vasuki and colleagues (2017) similarly labeled waveforms in accordance with their latency, utilizing names such as P100 (more commonly P1), N200 (more commonly P2), and N300 (more commonly P3). Dittinger et al., (2017) also used the label N200 (more commonly N2) as well as

N400 when describing a component peaking between 300-640 ms. For the purpose of clarity, the standardized labeling of components seen in Table 4 will be utilized when discussing data from all sources when suitable.

Dittinger et al., (2017) found significantly larger N2 and N400 amplitudes in the music training group compared to a control group measured during novel word learning, matching, and semantic tasks; these amplitude changes were attributed to faster and more efficient temporal processing. Similarly, Moreno and colleagues (2008) found significant enhancement of N2 amplitude to small pitch variations after 9 months of musical training, measured during music and speech tasks; these enhancements were attributed to more efficient pitch/frequency processing that can be generalized from music to speech perception.

Vasuki et al., (2017) utilized late AEPs to compare auditory and visual statistical learning in musically trained and untrained children; statistical learning was defined as the cognitive process of language acquisition through the identification of word boundaries. Recordings were obtained during familiarization phases of auditory and visual statistical learning triplet tasks (six different sets of three stimuli presented in succession). A statistically significant larger N2 (in response to the initial stimulus than in the final stimulus of a triplet set) was found in the musician group and not in the non-musician group for the auditory task; these findings were considered to show musicians' enhanced ability to extract statistical cues in speech, which is reported to be linked to reading, syntax comprehension, second language acquisition, and speech processing in adverse listening conditions.

Habili et al., (2016) longitudinally compared pitch perception and discrimination children with and without musical training. After two years of musical training, a decrease in P1 amplitude and an increase in an identifiable N1 (with an increase in N1/P1 ratio) was found when compared to baseline measurements in the passive pitch perception task; such changes were not found in the comparison groups. Decreased P1 latency was also seen in the music group compared to the sports (control) groups. Changes in the P1 component were considered to indicate a faster than average maturation in neural transmission in children with musical training. In the active pitch discrimination task, a larger P3 amplitude was reported in the music group compared to control groups; this finding was reported to reflect improved ability to detect deviations in pitch and may reflect enhanced auditory working memory.

Zuk and colleagues (2014) measured executive functioning in children with and without musical training utilizing function magnetic resonance imaging (fMRI). Participants completed an auditory-based set-shifting task during the fMRI recording. The musically trained children showed enhanced brain activation in regions associated with executive functioning (bilateral supplementary motor area and left ventrolateral prefrontal cortex) when compared to musically untrained children. These findings were said to support that musical training promotes the development and maintenance of executive functioning in children.

Behavioral Outcome Measures

Dittinger and colleagues (2017) utilized behavioral outcome measures to determine if musical training influenced word learning in children. Measures included phonological categorization tasks utilizing voicing, aspiration, and vowel length contrasts, a matching task in which participants had to match auditory and visual information learned in the pre-test phase, and a semantic task in which participants had to determine if a novel visual stimulus was semantically related to auditory information learned in the pre-test phase. Results indicated that musicians made significantly fewer errors than non-musicians on the phonological categorization tasks, matching task, and semantic task. Based on these results, musical training was considered to be associated with more efficient word learning.

Gerry and colleagues (2012) set out to determine if active engagement in music classes in infancy enhanced musical, communicative, and social development more than involvement in passive music classes. Behavioral measures included a sensitivity to Western tonality task in which infant's tonal preference (tonal or atonal) was measured via head-turn, a social-emotional development questionnaire (four subscales of the Infant Behavior Questionnaire), and an early communicative development questionnaire (Macarthur-Bates Communicative Development Inventories). On the tonal task, the active musical training group looked significantly longer to hear the tonal stimuli than the passive musical training group and the no musical training group; researchers considered this to indicate knowledge of Western tonality in only the active music group. When pre and post musical training scores were compared between active and passive musical training groups on the social-emotional development questionnaire, the musical training group showed significantly lower levels of distress to novel stimuli, more smiling and laughter, and were more easily soothed after 6 months of musical training; no significant differences were found pre-training. Significant differences in early communicative development were also reported after 6 months of musical training, with the active musical training group utilizing more gestures than the passive group, however, this result should be considered with caution as significant group differences were found pre-training. Based on these results, active musical training in infancy was considered to positively impact communication and social interaction as well as enhance culturally-relevant musical knowledge.

Habibi and colleagues (2016) utilized a behavioral tonal/rhythm discrimination task to determine the impact of two years of musical training on auditory processing in children. The tasks involved participants determining if a pair of short melodies were the same or different; two different sets of tasks were used, a tonal task in which pitch was altered, and a rhythm task in which duration was altered. When compared to groups with no training and sports training groups, the musically trained children more accurately detected differences in the tonal task; this was considered to indicate improved abilities to detect deviations in pitch in musically trained children.

Moreno and colleagues (2009) set out to determine if 6 months of musical training improved language-based processing utilizing behavioral pitch discrimination tasks and neuropsychological tasks and if performance differences between musician and non-musician children can be attributed to brain plasticity or a specific predisposition for music. The neuropsychological assessment included the IQ full scale of the Wechsler Intelligence Scale for Children (WISC-III) and a reading skills task in which participants were instructed to read aloud a set of words that included different grapheme-to-phoneme correspondences of increasing difficulty: simple and consistent, complex but consistent, and complex and inconsistent. The pitch discrimination tasks included a music-based task and a speech-based task where participants were instructed to decide if the last note/word in a set was normal or strange; participants were presented with sets of melodies and sentences in which the pitch of the final note/word was increased by either 35% (a weak incongruity) or 120% (a strong incongruity). Differences in scores on the WISC-III did not show significant differences between the musical training group and painting (control) group, however, performance on the reading task showed significantly lower errors for complex and inconsistent words in the musical training group. The performance similarities found between groups on the WISC-III and the differences found between groups on the reading task was considered to indicate that reading performance difference can be attributed to musical training effects and not maturation or a predisposition for music (as children were pseudo-randomly assigned into groups). Performance on the pitch discrimination tasks was significantly different between groups for the speech task but not the music task, with the musical training group making significantly less errors on weak incongruities on the speech task after 6 months of musical training. Although in the music task, the Group by Session and/or Congruity interactions were not significant, t-tests showed a trend of significantly less errors on weak incongruities in the music group. Improved performance on the pitch tasks was said to indicate improved discrimination of small variations in pitch as a result of musical training.

Slater and colleagues (2015) utilized the behavioral Hearing in Noise Test (HINT) to determine the impact of 1 versus 2 years of musical training on speech-in-noise perception. The standard HINT protocol was utilized in this study, with the children repeating English sentences presented via headphones in the presence of background noise presented at various signal-to-noise ratios (SNRs) to obtain an individual threshold SNR. A significant difference in HINT scores was found between groups. There was a significant improvement in HINT scores when pre and post training scores in the 2 years of training group were compared. Additionally, the 2 years of training group significantly outperformed the 1 year of training group. A correlation was also found, with better HINT performance linked to more hours of musical training. The researchers related their findings to previous research that suggests a 1dB improvement in SNR threshold can equate to a 10-15% improvement in speech recognition abilities. At least a 1 dB improvement was seen in 37% of children after 1 year of musical training and in 63% of the children after 2 years of musical training. Improved HINT scores after 2 years of musical training was said to indicate improved speech in noise processing in children.

Strait and colleagues (2012) utilized a similar research design to Slater and colleagues (2015) to determine the impact of musical training on speech-in-noise processing in children. In the present study, the aforementioned HINT test as well as the Words in Noise Test (WIN) were utilized as behavioral measures of speech-in-noise processing; standardized tests of attention and working memory were also utilized. The WIN protocol involved the children repeating words presented via a speaker at various SNRs to obtain an SNR threshold. The HINT protocol utilized in this study differed slightly from the aforementioned study, the location of the noise varied from 0° azimuth, -90° azimuth, and +90° azimuth and a composite SNR threshold score was then calculated. The Integrated Visual and Auditory Plus Continuous Performance Test was utilized to assess auditory and visual attention; the standard testing procedure was administered via a laptop, with automatically calculated scores generated. Auditory and visual working memory were also assessed via standardized tests: the Woodcock Johnson III Test of Cognitive Abilities Auditory Working Memory subtest, and the Colorado Assessment Tests 1.2 Visual Working Memory subtest. Significantly better HINT scores were seen in musicians compared to non-musicians when the speech and noise were spatially separated. Musicians also demonstrated better auditory (but not visual) working memory when compared to non-musicians. The composite results were said to indicate improved speech in noise processing in children with musical training through a top-down strengthening of cognitive abilities.

Strait and colleagues (2014) later set out to determine the impact of musical training on the perception of acoustically similar speech sounds in preschool-aged and school-aged children. Behavioral measures utilized included measures of IQ, working memory, and attention. The preschool-aged participants' IQ was measured using the Peabody Picture Vocabulary Test and the school-aged participants' IQ was measured using the Wechsler Abbreviated Scale of

Intelligence. Working memory was only analyzed in the school-aged participants; the Woodcock Johnson III Test of Cognitive Abilities Auditory Working Memory subtest and the Colorado Assessment Tests 1.2 Visual Span subtests were utilized. Attention was also only analyzed in the school-aged participants; the IHR Multicentre Battery for Auditory Processing attention subtests were utilized. School-aged musicians performed significantly better on the auditory working memory and auditory attention tasks compared to non-musicians, but not on the visual auditory working memory and visual attention tasks. These results were said to indicate enhanced auditory-specific cognitive abilities as a result of musical training through top-down strengthening of auditory processing.

Vasuki and colleagues (2017) utilized behavioral measures to determine the impact of musical training in childhood on auditory processing. Behavioral tasks were broken up into the following categories: auditory processing, cognitive processing, statistical learning. Auditory processing tasks included a musical skills task (Musical Ear test, melody and rhythm subtests), a frequency discrimination task (three-alternative forced choice trials, with a 1000 Hz tone as the standard), and the Dichotic Digits test. Cognitive processing tasks included a memory task (Clinical Evaluation of Language Fundamentals, forward and backward digit span subtests), a non-verbal intelligence test (Test of Non-verbal Intelligence), and a sustained attention task (Integrated Visual and Auditory Continuous Performance Task. Statistical learning tasks included auditory and visual statistical learning; the testing procedure was previously discussed in the objective outcome measures section of this systematic review as the same procedure was used for both behavioral and objective outcome measures. No significant differences between the groups were seen on the cognitive processing tasks. Musicians significantly outperformed non-musicians on the following auditory processing tasks: melody discrimination, rhythm discrimination, music score, and frequency discrimination. Additionally, musicians significantly outperformed non-musicians on the auditory statistical learning task, but not the visual statistical learning task. Correlational analysis also showed a moderate positive correlation between auditory statistical learning scores and music scores. Composite results were said to indicate improved auditory processing and statistical learning capacities in children who have undergone musical training.

Zuk and colleagues (2014) set out to determine the relationship between musical training and executive functioning in children. Behavioral measures utilized in this study include standardized cognitive assessments and a set-shifting task. Cognitive assessment included the following Delis-Kaplan Executive Function System subtests: Trail Making, Verbal Fluency, and Color-Word Interference. For the set-shifting task, children were presented with rules (either univalent or bivalent) via a visual cue and auditory stimuli indicating a left or right button press. Musically trained children performed significantly better than their untrained counterparts on the Coding, Verbal Fluency, and Trail Making cognitive assessments. No significant differences between groups were found for the set-shifting task.

DISCUSSION

The purpose of this systematic review was to determine the impact of childhood musical experience on speech perception and processing abilities. Studies included in this systematic review examined a variety of behavioral and objective outcome measures. Significant findings were present across all outcome measures, including MMN, cABR, P1-N1-P2 Complex, P3/P300, speech-in-noise testing, speech discrimination, auditory working memory/attention, reading, communicative development, and executive functioning tasks. All studies included in this systematic review controlled for confounding variables including previous musical training, socioeconomic status, and known hearing or neurological impairments.

Objective Outcome Measures

As seen in Table 5, all 7 of the 13 studies utilizing objective outcome measures found significant changes in amplitude of the respective AEP components as a result of musical training (Chobert et al., 2012; Putkinen et al., 2013; Putkinen et al., 2014; Dittinger et al., 2017; Moreno et al., 2009; Vasuki et al., 2017; Habibi et al., 2016). Increases in amplitude are representative of increased neural activation/processing in the associated areas of the cerebral cortex/brainstem (see Table 4). While amplitude can be associated with the amount of neural activation in a specific region or synchrony of activation, latency can be associated with the speed of neural response timing/processing. Three studies utilized latency measurements and found significant decreases in latency in AEP components as a result of musical training (Kraus et al., 2014a; Strait et al., 2012; Strait et al., 2013).

A cross-phaseogram procedure also quantifies response timing but utilizes a time-frequency matrix allowing for easy visual inspection; three studies utilized a cross-phaseogram procedure (Kraus et al., 2014b; Strait et al., 2014; Tierney et al., 2013). These three studies found significant improvements in neurophysiological distinction of the speech stimuli, indicated by phase shifts in the cross-phaseograms. To determine the neural encoding of a speech spectrum, a fast Fourier transformation can be utilized. Two studies utilized this procedure, finding more robust neural representation of harmonics as a result of musical training (Kraus et al., 2014a; Strait et al., 2012). In sum, all of the 13 studies utilizing objective outcome measures included in this systematic review (with 12 of 13 of the studies utilizing various AEP measurements and 1 of 13 (Zuk et al., 2014) utilizing fMRI) reported improved neurophysiologic processing of speech stimuli.

Behavioral Outcome Measures

Similar to the objective outcome measures utilized in the included studies, the behavioral outcome measures varied between studies. Of the 9 studies with behavioral outcome measures, 2 focused on measures of auditory processing (Habibi et al., 2016; Vasuki et al., 2017). When comparing performance on auditory processing tasks (see Table 2 for a description of tasks), musically trained children outperformed their untrained counterparts in pitch discrimination (Habibi et al., 2016) and melodic and rhythm discrimination (Vasuki et al., 2017) tasks. Vasuki and colleagues (2017) also found better auditory statistical learning abilities, defined as the ability to discriminate word boundaries by analyzing statistical relationships between syllables in a continuous stream of speech, in children with musical training.

Two studies utilized measures of speech-in-noise perception to determine the impact of childhood musical training on speech perception/processing (Slater et al., 2015; Strait et al., 2012). Slater and colleagues (2015) provided longitudinal evidence of improved speech-in-noise performance after 2 years of musical training and a relationship between total hours of training and performance, with more hours linked to better speech-in-noise perception. Similarly, Strait and colleagues (2012) found that musically trained children performed better than their untrained counterparts on the same measure of speech-in-noise perception utilized in the previously mentioned study (the HINT). The researchers also reported better performance on a standardized test of auditory working memory in the musically trained group (Strait et al., 2012).

Three studies included in this systematic review focused on measures of language development and linguistic ability to determine speech perception/processing abilities (Dittinger et al., 2017; Moreno et al., 2009; Strait et al., 2014). Dittinger and colleagues (2017) reported more efficient novel word learning for musically trained children compared to their untrained counterparts. Moreno and colleagues (2009) provided evidence of improved reading ability and pitch discrimination in children after 9 months of musical training; such improvements were not seen for children assigned to painting training. Strait and colleagues (2014) also reported better performance on tasks of auditory working memory and auditory attention in musician children compared to their untrained counterparts.

The final 2 studies utilizing behavioral outcome measures analyzed executive functioning abilities and global development (Gerry et al., 2012; Zuk et al., 2014). Gerry and colleagues (2012) found that infants who completed 6 months of active music training showed a greater increase in the use of prelinguistic communicative gestures, social-emotional development, and culture-specific musical knowledge when compared to infants who completed 6 months of passive music training (passive listening) and infants with no musical training. Zuk and colleagues (2014) reported that musically trained children outperformed their untrained counterparts on standardized measures of executive functioning focusing on processing speed, fluency, and cognitive flexibility (see Tables 2 and 5 for more information on measures used and significant findings).

Comparing Objective and Behavioral Findings

Importantly, 7 of the 16 studies included in this systematic review utilized both behavioral and objective outcome measures to determine the impact of childhood musical training on speech perception/processing. Dittinger and colleagues (2017) reported more efficient novel word learning reflected by behavioral performance on matching and semantic tasks and by the differences in amplitude of the electrophysiological data recorded during the behavioral tasks. Habibi and colleagues (2016) reported accelerated auditory processing via behavioral measures of pitch discrimination, electrophysiological data recorded during the pitch discrimination task, and a passive tonal perception task during which ERPs were measured. Similarly, Moreno and colleagues (2009) found both behavioral and objective correlates of enhanced pitch discrimination when electrophysiological data was recorded during behavioral tasks.

Strait and colleagues (2012) reported improved speech-in-noise perception reflected by behavioral and objective measures; behavioral speech-in-noise perception correlated with cABR response characteristics of spectral encoding of the speech stimulus and response latency. A similar study by the same researcher reported that behavioral performance on tasks of auditory working memory and attention were correlated with more distinct neural encoding of speech measured via cABR (Strait et al., 2014). Additionally, both Vasuki and colleagues (2017) and Zuk and colleagues (2014) reported behavioral and neural correlates of improved auditory statistical learning and executive functioning, respectively.

The above studies analyzing speech perception/processing via both objective and behavioral outcome measures provides valuable information on the relationship between such measures. All 7 of the above studies reported similar findings when the results of the behavioral and objective outcome measures were compared, implying that musical training impacts scores on both objective and behavioral tests of speech perception/processing in children in a similar fashion. Furthermore, these findings connect the musical training effect of enhanced neural encoding of auditory stimuli to real-world improvements in speech perception/processing in children.

Limitations

Considerable variability in musical training is noted upon analysis of the included studies. Half of the studies implemented a standardized musical training program while the remaining half employed questionnaires to limit the variance between length and rigor of the musical training between participants (see Table 3), thus justifying the need for future research implementing a standardized musical training program to control extraneous variance. Additionally, only 5 (31.25%) of studies employed a quasi-experimental design whereby subjects were randomly assigned into groups dictating their musical training (Chobert et al., 2012; Gerry et al., 2012; Kraus et al., 2014a; Kraus et al., 2014b; Moreno et al., 2009). As previously mentioned, musicians have been shown to have structural brain changes (Munte, Altenmuller, & Jancke, 2002; Schlaug et al., 1995; Moreno, 2009). Although studies have supported the concept that these structural and functional changes are due to musical training, not due to innate brain differences resulting in an increased likelihood for people to become musicians, a predisposition for such structural and functional brain changes cannot be ruled out at this time (Schneider et al., 2002; Hyde et al., 2009; Schulz et al., 2003). Therefore, future research is needed utilizing a quasi-experimental or randomized clinical trial design to further generalize results to a larger population and to rule out any predispositions for the aforementioned structural and functional brain changes influencing results.

As the impact of musical training on speech perception/processing has only recently gained attention by the audiology community, the pool of research currently available is limited. Half of the studies included in this systematic review were carried out at Northwestern University's Auditory Neuroscience Laboratory, justifying the need for additional studies or replications of the above studies by other researchers (Kraus et al., 2014a; Kraus et al., 2014b; Slater et al., 2015; Strait et al., 2012; Strait et al., 2013; Strait et al., 2014; Tierney et al., 2013).

CONCLUSION

The positive effect of childhood musical experience on speech perception and processing abilities is present throughout the literature reviewed when both objective and behavioral outcome measures are utilized. Additionally, studies that utilized both behavioral and objective outcome measures reported similar findings between the two methods. As a result, formal musical training in childhood should be considered as a viable option for auditory training when the goal is improved speech perception and/or processing. The results of these studies should also support the benefit of music classes in school curriculums to help children overcome communication challenges (such as listening in the presence of noise, distance, and poor acoustics) that are frequently found inside and outside of the classroom. Future research should address the limitations of the included studies, such as utilizing a standard musical training program and replicating the large proportion of research on this topic that originated from the Northwestern University Auditory Neuroscience Laboratory.

REFERENCES

Chobert, J., François, C., Velay, J. L., & Besson, M. (2014). Twelve months of active musical training in 8-to 10-year-old children enhances the preattentive processing of syllabic duration and voice onset time. *Cerebral Cortex*, *24*(4), 956-967.

Dittinger, E., Chobert, J., Ziegler, J. C., & Besson, M. (2017). Fast brain plasticity during word learning in musically-trained children. *Frontiers in human neuroscience*, *11*, 233.

Gerry, D., Unrau, A., & Trainor, L. J. (2012). Active music classes in infancy enhance musical, communicative and social development. *Developmental science*, *15*(3), 398-407.

Habibi, A., Cahn, B. R., Damasio, A., & Damasio, H. (2016). Neural correlates of accelerated auditory processing in children engaged in music training. *Developmental cognitive neuroscience*, *21*, 1-14.

Hannon, E. E., & Trainor, L. J. (2007). Music acquisition: effects of enculturation and formal training on development. *Trends in cognitive sciences*, *11*(11), 466-472.

Hübener, M., & Bonhoeffer, T. (2014). Neuronal plasticity: beyond the critical period. *Cell*, *159*(4), 727-737.

Hyde, K. L., Lerch, J., Norton, A., Forgeard, M., Winner, E., Evans, A. C., & Schlaug, G. (2009). Musical training shapes structural brain development. *Journal of Neuroscience*, *29*(10), 3019-3025.

Kishon-Rabin, L., Amir, O., Vexler, Y., & Zaltz, Y. (2001). Pitch discrimination: Are professional musicians better than non-musicians?. *Journal of basic and clinical physiology and pharmacology*, *12*(2), 125-144.

Kraus, N., Slater, J., Thompson, E. C., Hornickel, J., Strait, D. L., Nicol, T., & White-Schwoch, T. (2014). Auditory learning through active engagement with sound: biological impact of community music lessons in at-risk children. *Frontiers in neuroscience, 8*, 351.

Kraus, N., Slater, J., Thompson, E. C., Hornickel, J., Strait, D. L., Nicol, T., & White-Schwoch, T. (2014). Music enrichment programs improve the neural encoding of speech in at-risk children. *Journal of Neuroscience*, *34*(36), 11913-11918.

Liang, C., Earl, B., Thompson, I., Whitaker, K., Cahn, S., Xiang, J., ... & Zhang, F. (2016). Musicians are better than non-musicians in frequency change detection: behavioral and electrophysiological evidence. *Frontiers in neuroscience*, *10*, 464.

Martin, B. A., Tremblay, K. L., & Korczak, P. (2008). Speech evoked potentials: from the laboratory to the clinic. *Ear and hearing*, *29*(3), 285-313.

Moreno, S., Marques, C., Santos, A., Santos, M., Castro, S. L., & Besson, M. (2009). Musical training influences linguistic abilities in 8-year-old children: more evidence for brain plasticity. *Cerebral Cortex*, *19*(3), 712-723.

Moreno, S., & Besson, M. (2005). Influence of musical training on pitch processing: event-related brain potential studies of adults and children. *Annals of the New York Academy of Sciences*, *1060*(1), 93-97.

Moreno, S., & Bidelman, G. M. (2014). Examining neural plasticity and cognitive benefit through the unique lens of musical training. *Hearing research*, *308*, 84-97.

Micheyl, C., Delhommeau, K., Perrot, X., & Oxenham, A. J. (2006). Influence of musical and psychoacoustical training on pitch discrimination. *Hearing research*, *219*(1-2), 36-47.

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Annals of internal medicine*, *151*(4), 264-269.

Münte, T. F., Altenmüller, E., & Jäncke, L. (2002). The musician's brain as a model of neuroplasticity. *Nature Reviews Neuroscience*, *3*(6), 473-478.

Orduña, I., Liu, E. H., Church, B. A., Eddins, A. C., & Mercado III, E. (2012). Evoked-potential changes following discrimination learning involving complex sounds. *Clinical neurophysiology*, *123*(4), 711-719.

Pantev, C., Dinnesen, A., Ross, B., Wollbrink, A., & Knief, A. (2006). Dynamics of auditory plasticity after cochlear implantation: a longitudinal study. *Cerebral cortex*, *16*(1), 31-36.

Patel, A. D. (2011). Why would musical training benefit the neural encoding of speech? The OPERA hypothesis. *Frontiers in psychology*, *2*, 142.

Patel, A. D. (2014). Can nonlinguistic musical training change the way the brain processes speech? The expanded OPERA hypothesis. *Hearing research*, *308*, 98-108.

Picton, T. W. (2010). *Human auditory evoked potentials*. Plural Publishing.

Putkinen, V., Tervaniemi, M., Saarikivi, K., Ojala, P., & Huotilainen, M. (2014). Enhanced development of auditory change detection in musically trained school-aged children: a longitudinal event-related potential study. *Developmental science*, *17*(2), 282-297.

Putkinen, V., Tervaniemi, M., Saarikivi, K., de Vent, N., & Huotilainen, M. (2014). Investigating the effects of musical training on functional brain development with a novel Melodic MMN paradigm. *Neurobiology of learning and memory*, *110*, 8-15.

Rammsayer, T., & Altenmüller, E. (2006). Temporal information processing in musicians and non-musicians. *Music Perception: An Interdisciplinary Journal*, *24*(1), 37-48.

Schlaug, G., Jancke, L., Huang, Y., & Steinmetz, H. (1995). In vivo evidence of structural brain asymmetry in musicians. *Science*, *267*(5198), 699-701.

Schneider, P., Scherg, M., Dosch, H. G., Specht, H. J., Gutschalk, A., & Rupp, A. (2002). Morphology of Heschl's gyrus reflects enhanced activation in the auditory cortex of musicians. *Nature neuroscience*, *5*(7), 688-694.

Schulz, M., Ross, B., & Pantev, C. (2003). Evidence for training-induced crossmodal reorganization of cortical functions in trumpet players. *Neuroreport*, *14*(1), 157-161.

Slater, J., Skoe, E., Strait, D. L., O'Connell, S., Thompson, E., & Kraus, N. (2015). Music training improves speech-in-noise perception: Longitudinal evidence from a community-based music program. *Behavioural brain research*, *291*, 244-252.

Spiegel, M. F., & Watson, C. S. (1984). Performance on frequency-discrimination tasks by musicians and non-musicians. *The Journal of the Acoustical Society of America*, *76*(6), 1690-1695.

Steinschneider, M., & Dunn, M. (2002) Electrophysiology in developmental neuropsychology. In Segalowitz, S., & Rapin, I. (Eds.), Handbook of neuropsychology (2nd ed., Vol. 8, Part 1, Chap. 5, pp. 91-146). Amsterdam: Elsevier

Strait, D. L., Parbery-Clark, A., Hittner, E., & Kraus, N. (2012). Musical training during early childhood enhances the neural encoding of speech in noise. *Brain and language*, *123*(3), 191-201.

Strait, D. L., Parbery-Clark, A., O'Connell, S., & Kraus, N. (2013). Biological impact of preschool music classes on processing speech in noise. *Developmental cognitive neuroscience*, *6*, 51-60.

Strait, D. L., O'Connell, S., Parbery-Clark, A., & Kraus, N. (2014). Musicians' enhanced neural differentiation of speech sounds arises early in life: developmental evidence from ages 3 to 30. *Cerebral Cortex*, *24*(9), 2512-2521.

Sussman, E. S., Chen, S., Sussman-Fort, J., & Dinces, E. (2014). The five myths of MMN: redefining how to use MMN in basic and clinical research. *Brain Topography*, *27*(4), 553-564.

Tierney, A., Krizman, J., Skoe, E., Johnston, K., & Kraus, N. (2013). High school music classes enhance the neural processing of speech. *Frontiers in psychology*, *4*, 855.

Tremblay, K., Kraus, N., McGee, T., Ponton, C., & Otis, B. (2001). Central auditory plasticity: changes in the N1-P2 complex after speech-sound training. *Ear and hearing*, *22*(2), 79-90.

Tremblay, K. L., Shahin, A. J., Picton, T., & Ross, B. (2009). Auditory training alters the physiological detection of stimulus-specific cues in humans. *Clinical Neurophysiology*, *120*(1), 128-135.

Tzounopoulos, T., & Kraus, N. (2009). Learning to encode timing: mechanisms of plasticity in the auditory brainstem. *Neuron*, *62*(4), 463-469.

Vasuki, P. R. M., Sharma, M., Ibrahim, R., & Arciuli, J. (2017). Statistical learning and auditory processing in children with music training: An ERP study. *Clinical Neurophysiology*, *128*(7), 1270-1281.

Zuk, J., Benjamin, C., Kenyon, A., & Gaab, N. (2014). Behavioral and neural correlates of executive functioning in musicians and non-musicians. *PloS one*.