

**Short-term musical intervention enhances infants' neural
processing of temporal structure in music and speech**

Tian Zhao

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2015

Reading Committee:
Patricia K. Kuhl, Chair
Lynne A. Werner
Adrian KC Lee
Steven M. Demorest

Program Authorized to Offer Degree:
Speech and Hearing Sciences

© Copyright 2015
Tian Zhao

Table of Content

Abstract -----	4
Acknowledgement -----	5
Figure Titles -----	6
Chapter I: Introduction -----	8
I: Music learning as a model	
1. Frequency information	
2. Temporal information	
3. Gaps in the literature	
II: Current study	
1. Research questions	
2. Experimental design	
3. Hypothesis	
Chapter II: Methods -----	21
I: Subjects	
II: Stimuli	
III: Equipment and procedures	
IV: Data analysis	
Chapter III: Results -----	34
I: Music condition	
II: Speech condition	
Chapter IV: Discussion -----	36
I: Summary	
II: Theoretical interpretation	
III: Application of music intervention	
IV: Conclusion	
References -----	47

University of Washington

Abstract

Short-term musical intervention enhances infants' neural processing of temporal structure in music and speech

Tian Zhao

Chair of the Supervisory Committee:
Patricia K. Kuhl, Professor
Department of Speech and Hearing Sciences

Early music training provides an excellent model for studying neural plasticity and understanding the relations between music and speech processing; and it has implications for speech and language development. Most of the recent experimental evidence comes from cross-sectional studies of individuals with different music learning backgrounds. Data from these studies suggest an experience-related enhancement for high-level music pitch processing (e.g. chord and melody processing), and generalization to pitch processing in speech. However, little is known about processing in the temporal domain, where critical information lies both in music and speech, especially at the level of temporal structure. In music, temporal structure (e.g., meter) groups beats into units to help generate differential expectations for strong and weak beats (marching vs. waltz). In speech, temporal structure carries critical information such as syllable structure. The current study aimed to elucidate the effect of a short-term music intervention on infants' neural processing of temporal structure in music and speech using a random-assignment approach. Forty-seven 9-month-olds were randomly assigned to complete 12 sessions (~4 wks) of social, multimodal music intervention (Intervention group) or 12 sessions of social free play (Control group). The mismatch response (MMR) was the target neural measure used to quantify infants' processing of temporal structure both in music and in speech after the intervention/control sessions. Using Magnetoencephalography (MEG) that has good spatial resolution, we compared the MMRs in the temporal cortical regions as well as the prefrontal regions between the Intervention and the Control groups. The results showed that the Intervention group exhibited larger MMRs in both temporal and prefrontal regions than their Control group counterparts, indicating stronger ability to track high-level temporal structure, as hypothesized. It is the first known experiment demonstrating an effect of music intervention in the temporal information processing domain in infancy, and the first to show generalization from music intervention to speech processing in infancy. The results provide important implications for our understanding of auditory experience-related effects on music and speech early in development.

Figure Titles

Figure 1

Hierarchical organization of temporal structure in music

Figure 2

Hierarchical organization of temporal structure in speech
(from Goswami & Leong, 2013)

Figure 3

a: Schematics of stimuli in the music condition

b: Schematics of stimuli in the speech condition

Figure 4

Inflated surface of the head template and the selection for bilateral auditory cortices (blue) and prefrontal regions (red)

Figure 5

a: Mismatch response from the auditory regions for the Intervention group (blue) and the Control group (green) in music condition

b: Mismatch response from the prefrontal regions for the Intervention group (red) and the Control group (orange) in music condition

Figure 6

a: Mismatch response from the auditory regions for the Intervention group (blue) and the Control group (green) in speech condition

b: Mismatch response from the prefrontal regions for the Intervention group (red) and the Control group (orange) in speech condition

Figure 7

Statistical comparison of MMR in music condition

Figure 8

Statistical comparison of MMR in speech condition

Figure 9

Schematic of the theoretical framework

Acknowledgement

My dissertation work would not have been possible without the help and support from many individuals. Over the past few years, I have grown in many ways through working on this project: to work with people from different backgrounds, to truly appreciate an interdisciplinary and collaborative environment and of course, to explore in research topics where I find my passion.

Dr. Patricia Kuhl has guided me in fundamental ways that I am extremely grateful for, both in theoretical exploration and in application. Her theoretical insights of exploring music intervention in infancy have laid the foundation of the current work. I have also learned a great amount from each of my doctoral committee members (Drs. Lynne Werner, Adrian KC Lee, Steven Demorest and Richard Wright), who are the leading experts in areas of psychoacoustics, phonetics, music cognition and MEG technology.

The support from I-LABS members is crucial in completing a large-scale experiment such as my dissertation project. From the scheduling and testing infants over a 6-week period of time by all the research assistants (special gratitude for Denise Padden, who is always calm and can think of all the details; as well as Nicole Miles, who completed a substantial amount of the work with me), to constant technical support by Justin Reichel and Cary Dubinsky, to all the MEG technology related help (special gratitude for Myles Reilly for his constant hard work in helping setting up all MEG recording sessions as well as Dr. Rey Ramirez in his willingness to teach the underlying mathematics in processing MEG data at a level that is most appropriate for my background as well as sharing the softwares).

The support from my family and my mentors has been most critical for me to be where I am today. Dr. Christine Moon, Dr. Dana Anderson and Dr. Gregory Youtz from Pacific Lutheran University, have provided tremendous guidance for me during my undergraduate studies, which paved my road in further pursuing my graduate studies. My parents and grandparents, who have always supported me in pursuing my own passion in knowledge and science, make me feel safe in this long journey of exploration.

Last but not least, this work cannot be done without the support of all the families who are devoted in supporting scientific research as well as the generous funding from National Science Foundation (NSF) LIFE center grant.

Chapter I: Introduction

I. Early music training as a model

Early music training has received increased attention in recent years as an effective model to investigate experience-related neural plasticity (Zatorre, 2013). It is generally defined as formal instruction that started at a young age (before age 7) and it involves complex skill learning over time, as music performance requires precise coordination of multiple sensory and motor systems. A large amount of training and practice is needed to reach a high level of proficiency.

Over the last decade, researchers have extensively examined two main areas regarding the neural plasticity related to early music training: 1) effects within the music domain, that is, the effects of early music training on the processing of music sounds (Herholz & Zatorre, 2012; Munte, Altenmuller, & Jancke, 2002; Schlaug, 2001); and 2) the generalization of effects from early music training to speech sound processing (Besson, Chobert, & Marie, 2011). Generalization effects to speech processing from music training are particularly interesting in terms of both theory and practice. Music and speech are both important and universal components of the human-generated acoustic environment (Brown, 1991). Both require rapid and accurate processing of complex frequency and temporal information, and the extraction of high-level abstract meanings from dynamic acoustic inputs. Researchers have been interested in the relations between music and speech for centuries and investigations have been conducted at many levels, from the encoding of sounds in the auditory system to cultural impacts at different developmental stages (McMullen & Saffran, 2004; Patel, 2008).

Theoretically, examining the generalization effects from early music training to speech sound processing may enrich our understanding of the shared systems that underlie music and speech processing as well as the neural mechanisms that support such cross-domain effects. In practice, speech processing is extremely important in human communication especially early in the development, since early speech processing is critically associated with later language ability (Kuhl et al., 2008). Studying generalization effects from early music training to speech processing may also provide potential alternative intervention methods for communication disorders.

1. Frequency information

The majority of current studies have focused on measuring the effects of early music training on the processing of frequency information; and the effect of music training has been commonly assessed by comparing a group of individuals who reportedly have received extensive formal music training early in life to a group of untrained individuals (i.e. cross-sectional). A series of studies within the music domain have suggested that adult musicians exhibit stronger representation of the pitch of complex musical tones, especially when the timbre of the complex tones matches their primary instruments (Pantev et al., 1998; Pantev, Roberts, Schulz, Engelien, & Ross, 2001; Schneider et al., 2002). Furthermore, adult musicians show better discrimination of small frequency differences, but only when the frequency difference is embedded in a familiar musical context, for example, a chord structure (Koelsch, Schroger, & Tervaniemi, 1999; Tervaniemi, Sannemann, Noyranen, Salonen, & Pihko, 2011), an interval, or a melody (Fujioka, Trainor, Ross, Kakigi, & Pantev, 2004, 2005).

These results have also been observed in longitudinal studies and laboratory training

studies in adults and children. Lappe and colleagues (2008) trained a group of previously untrained adults using a 2-week sensorimotor auditory music-training paradigm. After the training, participants exhibited better detection of pitch violations embedded in melodies than controls that experienced passive auditory training. Children who are involved in Suzuki music lessons were tracked over the course of one year in comparison to a group of control children and their neural responses were measured in response to auditory stimuli. Fujioka and colleagues (2006) observed that in 6 years-olds, the evoked neural response increased in amplitude and decreased in latency for the Suzuki children compared to their control counterparts, but only for the violin tones and not for broadband noise. Similarly, in 4-5 year-olds, Shahin and colleagues (2008) observed stronger induced cortical oscillation in the gamma range in the Suzuki group.

More interestingly, generalization effects from early music training to pitch processing in speech have been repeatedly demonstrated (Kraus & Chandrasekaran, 2010). Cross-sectional studies investigating adults and children have indicated that musically trained individuals exhibit: 1) better frequency information encoding at the subcortical level, and 2) better pitch processing at the cortical level, for both native and non-native speech sounds (Magne, Schon, & Besson, 2006; Marie, Delogu, Lampis, Belardinelli, & Besson, 2011; Moreno & Besson, 2006; Musacchia, Sams, Skoe, & Kraus, 2007; Parbery-Clark, Skoe, Lam, & Kraus, 2009; Parbery-Clark, Tierney, Strait, & Kraus, 2012; Schon, Magne, & Besson, 2004; Wong, Skoe, Russo, Dees, & Kraus, 2007). For example, Chandrasekaran and colleagues (2009) compared the cortical processing of foreign lexical tone differences of musicians and nonmusicians. An oddball paradigm was used to measure subjects' preattentive discrimination of pitch contours

representative of Mandarin lexical tones. English-speaking musicians exhibited a larger discriminative response to a change in pitch contour than the nonmusicians. Another study documented differences between adult musicians and nonmusicians in terms of the cortical and subcortical measures (Musacchia, Strait, & Kraus, 2008). Both cortical and subcortical responses to speech syllables were extracted from EEG recordings of participants' continuous neural activity. At the subcortical level, musicians demonstrated a better frequency following response to the fundamental frequency of the speech sounds. At the cortical level, musicians exhibited sharper timing to the speech sounds than their nonmusician counterparts. More importantly, the subcortical measure was significantly correlated with the cortical measure, and both measures correlated with years of music training in musicians. Researchers have argued for shared mechanisms in processing music and speech information, which may exert influence in both domains at an early stage of auditory processing.

2. Temporal information

Despite the critical role temporal information plays in music and speech, only a few studies have examined the effect of the early music training on the processing of temporal information.

Hierarchical organization

Temporal information is hierarchically organized in both music and speech (Essens, 1986; Martin, 1972; Povel & Essens, 1985). In the music domain, temporal information exists at multiple levels (Drake, Jones, & Baruch, 2000). For example, at the rhythm level, the note-to-note intervals are short and highly variable (Figure 1). However, the underlying isochronous beats with precise timing can be realized, even with the absence

of notes (beat level). Furthermore, beats are organized into units by contrasting the intensity level of the beats (strong vs. weak) and such units establish the metrical structure of music. Therefore, two pieces of timing information are critical in the perception of musical temporal structure: onset of a unit and the number of beats in one unit. For instance, marching music usually follows the duple meter structure: (2 beats per unit: strong-weak) whereas waltz follows the triple meter structure (3 beats per unit:

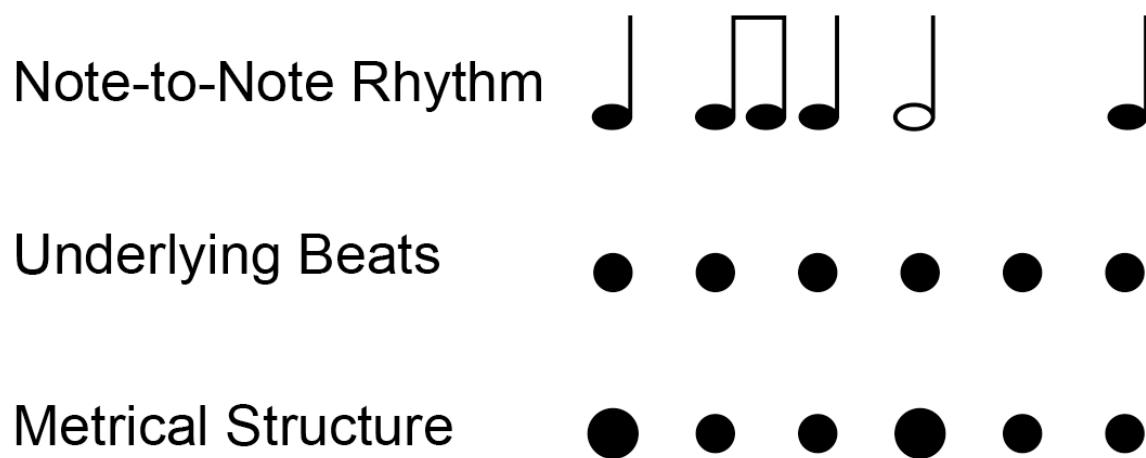


Figure 1 Example of a hierarchical organization of temporal structure in music. At the note-to-note level, the intervals between notes are variable but the isochronous beats that underlie these notes can be realized. These beats are further grouped into units of three through increasing the intensity of the first note. This grouping establishes a specific metrical structure (triple meter) where strong and weak beats are expected at specific time points.

strong-weak-weak). Therefore, the ability to accurately track the timing of the strong beat is critical in the processing of musical temporal structure (Honing, 2012).

In the speech domain, temporal information spans a wide range, from fine structure (600Hz to 10kHz) to the overall fluctuation, often termed amplitude modulation, or envelope (2-50Hz) (Goswami & Leong, 2013; Poeppel, Idsardi, & van Wassenhove, 2008; Rosen, 1992). Here, we will discuss temporal information contained in the envelope. In a recent model, Goswami and Leong (2013) provided a way to characterize

different rates of amplitude modulation in speech and the information contained at each level. Segmental information such as phonetic contrasts (e.g. voice vs. voiceless) is conveyed via the fastest rates of amplitude modulation (Figure 2C), followed by information on syllable structure at a slower rate of amplitude modulation (Figure 2B). Prosodic information (e.g. stress pattern) can be retrieved from the slowest rates of amplitude modulation (Figure 2A). Speech sounds remain generally intelligible with slower rates (<16 Hz) of amplitude modulation, even when the spectral information is highly degraded, demonstrating the important role of amplitude modulation, or envelope information, in speech perception (Shannon, Zeng, Kamath, Wygonski, & Ekelid, 1995). We consider to the slower rates of amplitude modulation as the temporal structural information in speech.

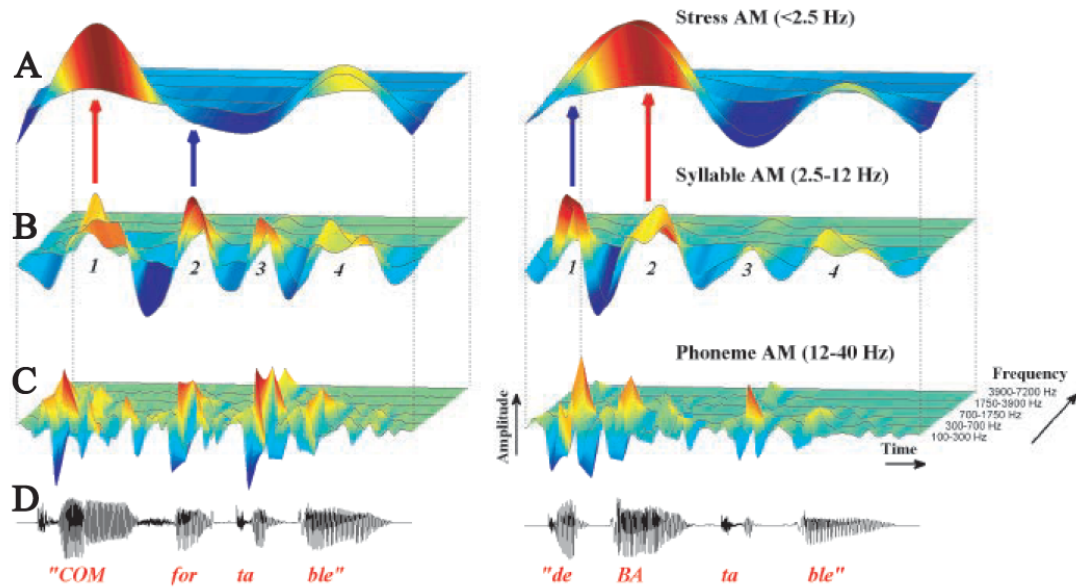


Figure 2 Hierarchical organization of temporal structure in speech, from Goswami & Leong (2013). In panel D, the waveforms are exhibited for two words with stressed syllables at different locations. In panel C, the fast envelope amplitude modulation (AM) is exhibited which carries information at the phoneme level. In panel B, envelope AM at a slower rate is exhibited which carries mostly the information of the timing of the syllables; and at the slowest envelope rate (panel A), the prosodic/stress pattern information can be extracted.

Previous evidence

Within the music domain, cross-sectional studies suggest that adult musicians are better than nonmusicians in terms of detecting violations in the high-level temporal structure in music (i.e. the metrical structure), but groups do not differ in detecting violations in the note-to-note rhythmic pattern (Geiser, Sandmann, Jancke, & Meyer, 2010; Vuust et al., 2005). This result is consistent with findings from the pitch processing domain in that differences between individuals with and without early music training lie in the processing of information embedded in a structural context (e.g. pitch violation in a tonal melody or a metrical structure). To date, only one study has examined generalization effects from early music training to speech processing in terms of temporal

information, and it demonstrated that adult musicians are more sensitive than nonmusicians to violations in syllable structure (Marie, Magne, & Besson, 2011).

3. Gaps in the literature

The literature to date prevents a strong causal conclusion about the effects of early training in the music domain, and the generalization of the effects from early music learning to the speech domain for several important reasons. First and foremost, while the cross-sectional approach to studying the differences between musically trained and untrained individuals is very informative, it does not rule out the potential that genetically based differences in auditory acuity contribute to the self-selection process which lead individuals with high acuity to musical training and others to no musical training. Second, the cross-sectional approach also does not allow the control of the onset of music training, the nature of the music learning experienced, or the amount of music intervention outside of the formal learning environment. Studies often use different arbitrary selection criteria to define ‘musically trained’ vs. ‘untrained’ individuals, yet each criterion may be an important factor for the outcome of training, and a better control of these factors will allow a better understanding of their contributions to early music learning (Schlaug, Jancke, Huang, Staiger, & Steinmetz, 1995; Wong, Roy, & Margulis, 2009). Third, most studies demonstrating generalization effects to speech processing did not include similar measurements in the music processing domains in the same experiment, making it difficult to argue for a shared system that underlies both processing domains. Lastly, as discussed earlier, to date there are only a few studies investigating the effects of early music training on temporal information processing.

To address these issues, the current study incorporated several important and novel techniques. First, early music training was operationalized as a music intervention procedure in a controlled laboratory setting for 9-month-old infants with no previous or concurrent enrollment in infant music classes. This approach allows maximal control of the music intervention: infants received the same amount of music intervention of the same nature across the same time period. This approach also minimizes effects from implicit learning due to intervention to music outside of the laboratory prior to the study (Bigand & Poulin-Charronnat, 2006). Second, infants were randomly assigned to be in the Intervention group or the Control group, minimizing the effects of genetic predisposition and allowing for strong conclusions that any observed differences between groups are due to the difference in the experience (Gerry, Unrau, & Trainor, 2012; Kuhl, Tsao, & Liu, 2003). Finally, infants' abilities to process temporal structure in both music and speech were measured in comparable paradigms, allowing for the investigation of effects of music intervention in the music domain, and generalization to speech domain in one study.

II. Current Study

1. Research questions

The current study aimed to provide information on the effect of music intervention at 9 months of age. Specifically, we focused on infants' neural processing of temporal structure in music and generalization to the processing of temporal structure in speech. Two main research questions were investigated: 1) Does a short-term music intervention enhance infants' neural processing of temporal structure in music (i.e., metrical structure)? 2) If so, does the effect generalize to the speech domain?

2. Experimental design

Nine-month-old infants were randomly assigned to an Intervention or a Control group. Infants in the Intervention group completed 12 sessions of social, multimodal music intervention over a 4-week period of time, learning triple meter structure. Infants in the Control group completed 12 sessions of social free play. Following the 12 sessions, infants' abilities to process temporal structure were quantified by the mismatch responses (MMRs), which measure sensitivity to violation of metrical structure in music and to violation of syllable structure in speech. MMRs were extracted from continuous recordings of infants' neural activities using Magnetoencephalography (MEG), which provides good spatial resolution, allowing the investigation of effects in different regions of the brain.

Intervention/Control design

The music intervention sessions targeted triple meter learning by synchronizing infants' movements to musical beats in a social environment. Several steps were taken to ensure maximal learning specific to the laboratory-controlled music intervention sessions.

First, triple meter was chosen as the target meter because: 1) it is uncommon (5-10% occurrence) in the Western music system (Fraisse, 1982), thus minimizing learning through exposure outside of the lab, and 2) it has been demonstrated that 9-month-olds are capable of detecting pattern change in duple meter, but not in triple meter (Bergeson & Trehub, 2006). Second, synchronizing movement to the beats has been shown to be an effective auditory training activity for both adults and infants. For example, adult nonmusicians improved their ability to process both frequency and temporal information in music only after a sensorimotor plus auditory training procedure (Lappe et al., 2008; Lappe, Trainor, Herholz, & Pantev, 2011). In addition, infants demonstrated preferential looking to a specific metrical structure after being bounced in a corresponding time pattern for 2 minutes (Phillips-Silver & Trainor, 2005, 2007). Third, a social environment is also a critical component for infants' learning. Previous work has suggested that infants learn foreign speech contrasts only from a live tutor, not from TV or audio only; and that they learn more from foreign language intervention on TV when a peer is present (Kuhl et al., 2003; Roseberry & Kuhl, 2013). Lastly, the Control group completed 12 sessions of social free play in the same room with the same experimenters. This allows control for effects related to routine exposure to this particular social environment. Thus, the difference observed between the Intervention and Control group can be attributed the experience of the music intervention.

MEG measurement

Infants completed two MEG test sessions that measured their sensitivity to a violation of triple meter structure (music condition) and a violation of syllable structure (speech condition) within 2 weeks of the last intervention/control session. The order of the

sessions was counterbalanced across all infants. In both the music and speech conditions, an oddball paradigm was used where a standard structure was established through repetitive standard trials (84%) and the structure was violated occasionally by deviant trials (16%). Continuous neural activity, reflected by the changing magnetic field, was recorded by the MEG. The mismatch response (MMR), defined as the difference wave between response to standard trials and deviant trials during a target time period, served as the dependent measure to quantify infants' ability to process temporal structural information in music and speech conditions.

The mismatch response (MMR)

The mismatch response was originally described in EEG literature as a negative peak between 150ms and 250ms in the difference wave between response to standard sounds and deviant sounds in adults, and interpreted as a reflection of sensitivities to the acoustic differences in the sounds (Näätänen, Paavilainen, Rinne, & Alho, 2007). It has been consistently observed in infancy and later developmental stages, and has been a useful tool in examining auditory system development and speech sound learning (Cheour et al., 1998; Cheour et al., 2004; Garcia-Sierra et al., 2011; Gomot, Giard, Roux, Barthelemy, & Bruneau, 2000; Imada et al., 2006; Kuhl, 1998; Kuhl et al., 2008; Rivera-Gaxiola, Klarman, Garcia-Sierra, & Kuhl, 2005; Rivera-Gaxiola, Silva-Pereyra, & Kuhl, 2005). Importantly, the MMR has been employed to study sensitivity to violations of structural regularities or predictions, suggesting the involvement of higher-level central processes reflected in the MMR response in adults (Bekinschtein et al., 2009; Bendixen, SanMiguel, & Schroger, 2012; Winkler, Denham, & Nelken, 2009), and very recently, in young infants (Basirat, Dehaene, & Dehaene-Lambertz, 2014).

Two regions have been identified as the primary underlying sources of the MMR: the bilateral auditory cortices (temporal regions) and prefrontal regions (Näätänen et al., 2007). Strong evidence from various methods in addition to EEG (e.g. MEG, animal models and lesion data) supports the contributions of the auditory cortices (Alho, 1995; Rosburg et al., 2005); however, evidence supporting the prefrontal regions is weaker, based on source modeling of EEG data (Giard, Perrin, Pernier, & Bouchet, 1990; Ortiz-Mantilla, Hamalainen, & Benasich, 2012). EEG has poorer spatial resolution than MEG, because the electrical field is dispersed when conducting through variable tissue layers before measurement (e.g. cerebral fluid, scalp etc.), while the magnetic field measured in MEG is not influenced by intervening tissue (Hamalainen, Hari, Ilmoniemi, Knuutila, & Lounasmaa, 1993). One recent study using MEG measurement in infants has indicated activation of the inferior frontal gyrus in the MMR in response to speech contrasts (Kuhl, Ramirez, Bosseler, Lin, & Imada, 2014). Therefore, a secondary question of the current study was to examine whether the effects of music intervention are observed in both temporal and prefrontal regions, taking advantage of the good spatial resolution of the MEG.

3. Hypotheses

To summarize, the current study examined the effect of a short-term music intervention at 9 months of age on infants' neural processing of temporal structure in both music and speech. Specifically, we hypothesized that the Intervention group would exhibit larger MMRs than the Control group in both conditions, and that the effects would be similar in the temporal and prefrontal brain regions.

Chapter II: Method

I. Subjects

Forty-seven infants born and raised in monolingual English-speaking families were recruited at 40 weeks of age. The inclusion criteria included: 1) full term and born within 14 days of due date, 2) no known health problems and no more than 3 ear infections, 3) birth weight in the range of 6-10 lbs., and 4) no previous or concurrent enrollment in infant music classes. All experimental procedures were approved by the Institute Review Board of the University of Washington.

Infants were randomly assigned to either the Intervention group (N=22) or the Control group (N=25). Reported music in their home environments was comparable between the two groups of infants (Intervention: 9.93 ± 6.83 hrs/week; Control: 12.89 ± 9.47 hrs/week). Participation required completion of 12 intervention or control sessions over a 4-week period, and 3 MEG recordings (1 for the music condition, and 2 for the speech conditions) on two separate days within 2 weeks of the last intervention or control session. Overall, 1 infant failed to complete all intervention sessions and 7 failed to complete MEG recording due to fussiness. In addition, 3 MEG recordings from the music condition and 8 from speech condition failed to produce usable data due to: excessive movement (2), too few usable trials (2) and technical failure (7). A total of 36 MEG recordings were included in analysis of the music condition (18 Intervention, 12 male; 18 Control, 9 male) and 35 sets of 2 MEG recordings were included in analysis of the speech condition (16 Intervention, 12 male; 19 Control, 9 male).

In addition, infants with successful MEG recordings were recruited to complete a structural MRI scan within 2 weeks of the last MEG recording. MRI scans from one subject were used to construct the head model.

II. Stimuli

1. Intervention phase

Intervention group

Recordings of children's songs in triple meter were selected from various commercially published music CDs for infants and toddlers. They were selected to cover a wide range of tempi (slow to fast) and voices (adult and children, male and female, single and multiple voices). Songs were recorded on 6 CDs of about 15 minutes in duration.

2. MEG testing phase

Music condition

A complex tone (Duration: 200ms, Sampling frequency: 44.1kHz) with fundamental frequency of 220 Hz (A3) was synthesized by combining a tone with 'grand piano' timbre with a woodblock sound in Overture software (Version 4, Sonic Scores). This tone served as a 'weak' tone. A 'strong' tone was created by amplifying the weak tone by 10dB in Audacity software (Version 2.0, Audacity(R) software is copyright (c), 1999-2012, Audacity Team.).

Speech condition

Disyllabic nonwords were created in Praat by repeating a synthesized syllable /bi/ (Boersma & Weenink, 2009). The syllable /bi/ was synthesized (Duration: 160ms, Sampling frequency: 44.1kHz and fundamental frequency: 220Hz) to have 30ms of

formant transition at the beginning and at the end as well as 100ms of steady state vowel. For the standards, a disyllabic nonword /bibbi/ was created by combining two syllables with 150ms of silence in between. For the deviants, /bibi/ was created by reducing the duration of the silence to 50ms. For both stimuli, the first syllable was amplified by 5dB to create a strong-weak stress pattern.

III. Equipment and procedures

1. Intervention phase

Intervention group

Infants assigned to the intervention group completed 12 sessions of social, multimodal music intervention over a 4-week period of time. All sessions were 15 minutes long. The sessions took place in a sound-attenuating booth decorated to be infant friendly. In each session, 1 of the 6 CDs was played through 2 speakers at a comfortable listening level of 65 dBA SPL, measured at the center of the room. Four video cameras were placed at different locations in the room to capture the behaviors of the infants during all sessions. Up to 3 infants and their primary caregivers were in the room, along with an experimenter who facilitated the session. The caregivers were instructed to interact with the infant throughout the sessions, with the aim of synchronizing the infants' movements to the musical beats. A variety of infant-safe simple percussive musical toys were provided to facilitate infants' movement, such as shaking maracas. Foot tapping and bouncing were also used. The caregivers were also instructed not to practice similar procedures at home.

Control group

Infants assigned to the Control group completed 12 sessions of social free-play. In each session, up to 3 infants and their primary caregivers were in the room, along with an experimenter. The infants engaged in free play in the room with their caregivers, the other infants as well as the experimenter. Non-musical toys appropriate to the infants' age were provided to facilitate the social play.

2. MEG testing phase

Infants completed 3 MEG recordings on two separate days within 2 weeks of the last intervention/control session. One recording for the music condition and two recordings for the speech conditions were made and the order of the recordings was counterbalanced across infants.

Stimulus presentation

Stimuli were delivered using the Psychophysics Toolbox in MATLAB (Brainard, 1997) on a HP workstation connected to TDT RP 2.7 Hardware (Tucker-Davis Technologies hardware, Alachua, FL). All stimuli were processed such that their RMS values were referenced to 0.01, and they were further resampled to 24414Hz for the TDT. Subsequently, the sounds were played through a speaker with a flat frequency response at a comfortable listening level of 65 dB SPL, measured under the MEG dewar.

In the music condition, a continuous triple meter structure was created by combining a strong tone with two weak tones with sound-onset-asynchrony (SOA) of 300ms. Occasionally, this pattern was violated by removing the last 'weak' tone. The strong tones following the violation are considered deviants, while the strong tones prior to the violation are considered standards (Figure 3a). There were 1250 trials in total with 200 deviant trials (16%).

In the speech condition, two separate stimulus sequences were created to address the acoustic differences between the standard /bibbi/ and deviant /bibi/. In a long recording, 1250 trials were played of which 84% were standards /bibbi/ and 16% were deviants /bibi/. The SOAs was jittered between 900ms and 1100ms to minimize effects associated with predicting the onset of each nonword. In a short recording, 200 trials of deviants /bibi/ were played to establish the standard response to that stimulus. Therefore, the difference wave was defined as response to /bibi/ in the short recording subtracted from response to the same stimulus in the long recording (Figure 3b).

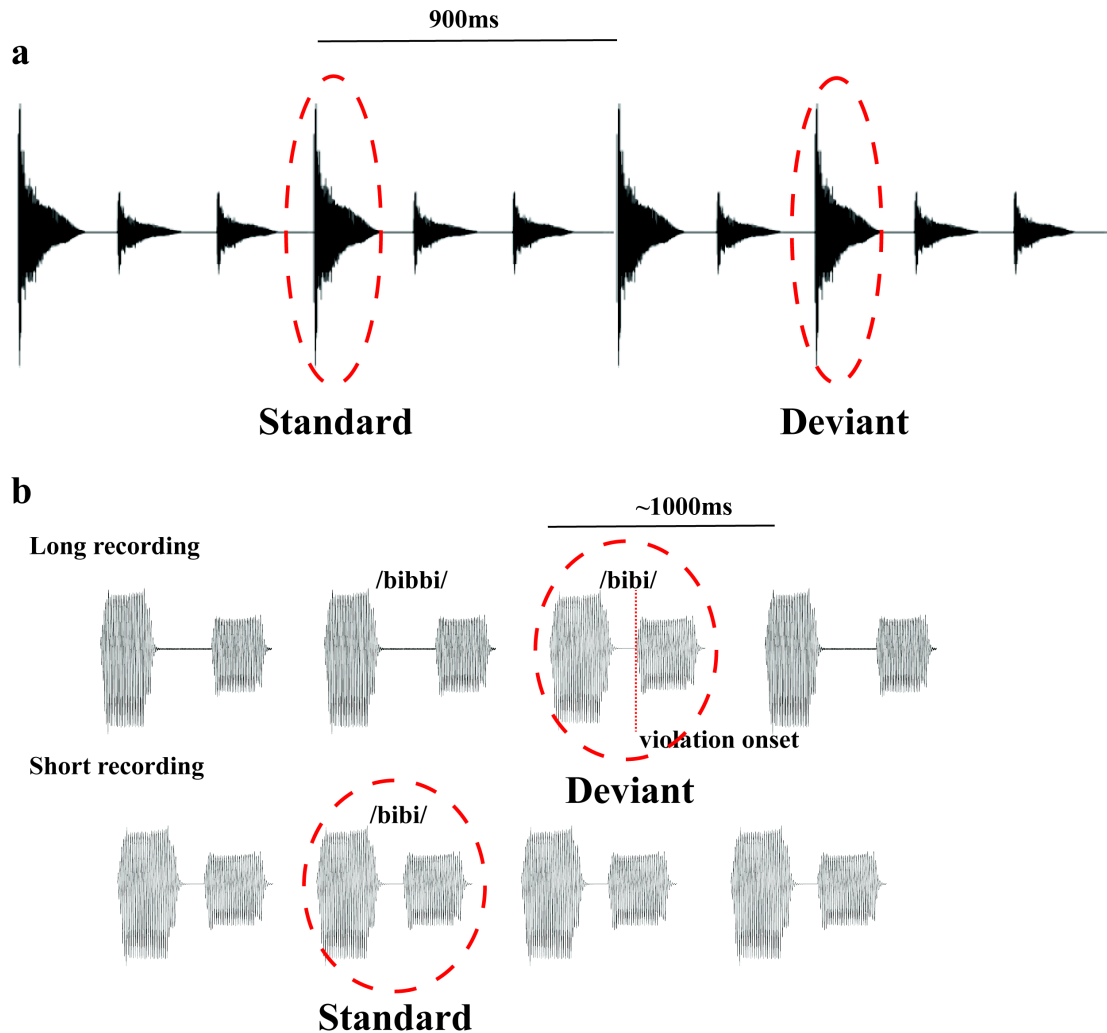


Figure 3 a) Schematics of stimuli in the music condition; while standard and deviant sounds are acoustically identical, deviants violate the standard temporal structure occasionally (16%). b) Schematics of stimuli in the speech condition; in the long recording, deviants /bibi/ violate the standard temporal structure occasionally (16%), however, the frequent /bibbi/ and deviant sounds /bibi/ are not acoustically identical. A short recording with only the deviant /bibi/ was played during a separate short MEG recording to establish the standard response to the deviant sound /bibi/. And the difference wave is therefore defined as the response to deviant /bibi/ in the long recording minus response to /bibi/ in the short recording. This comparison allows us to examine MMR related to the detection of temporal structure violation only. Note that the onset of violation is the onset of the second syllable (210ms after the onset of the nonword).

MEG measurement

All Magnetoencephalography (MEG) data were acquired inside a magnetically shielded room (MSR: IMEDCO) using a MEG (306-channel dc-SQUID VectorView) system (Elekta-Neuromag) with 204 planar gradiometers and 102 axial magnetometers. All data were acquired at a 1kHz sampling frequency.

In a typical MEG session, the infant was first seated in a customized high chair outside of the MSR. A research assistant distracted the infants while the technician fit a stretchy cap on infants' head. One binocular pair of electro-oculogram electrodes was attached to measure eye blinks (EOG). Five Head Position Indicator (HPI) coils were attached to the stretchy cap to measure head position continuously under the MEG dewar. Three landmarks (LPA, RPA and nasion) and the 5 HPI coils were digitized along with 100 additional points along the head surface with an electromagnetic 3D digitizer (Fastrak®, Polhemus, Vermont, U.S.A). Then the infant was placed under the MEG dewar, sitting on a similar customized chair. A research assistant continued to distract the infant with toys, while the primary caregiver was seated next to the MEG machine. Once the infant appeared to be calm and alert, the MEG recording started and the stimulus presentation began. Throughout the MEG recording, all individuals in the MSR except for the infants wore earplugs that significantly attenuated the sounds to minimize potential influence of the adults reacting to the sounds. The recording ended when all trials were presented or when the infant became too fussy to continue.

In addition, at the end of each MEG session, a 5-minute empty-room recording was made with the same stimuli playing.

3. MRI Structural Scan

The MRI structural scans were completed within 2 weeks after the last MEG session using a 3.0T system with an 8-channel head coil (Achieva, Phillips, Netherlands). A multi-echo T1 pulse sequence (3D water excited/TFE) was used with the following parameters: TR =24ms, TI=1450ms and TEs=6.5ms, 12.2ms &18ms; acquisition voxel size 0.37 mm³, SENSE factor 2.5 in the anterior-posterior direction.

IV. Data Analysis

1. Template Creation

MRI scans obtained from one participant employing the abovementioned procedures were used to create the template head model. The images were first processed by calculating the root-mean-square (RMS) of the values obtained from the 3 echoes for each voxel. The resulting images were segmented in FSL-FAST (Zhang, Brady, & Smith, 2001). The white matter component resulting from the segmentation was then used to process the images again in order to enhance the signal for the white matter. Cortical reconstruction and volumetric segmentation were performed using the FreeSurfer image analysis suite (<http://surfer.nmr.mgh.harvard.edu>). A surface based cortical source space was created using the topology of a recursively subdivided icosahedron 5, resulting in approximately 20,484 source points distributed throughout cortical surfaces. In addition, a subcortical volumetric source space with grid spacing of 5 mm was constructed, including approximately 4,425 source points distributed throughout subcortical structures and the cerebellum.

2. Preprocessing

The raw MEG recordings underwent a series of standardized preprocessing steps to

suppress noise. The temporal signal space separation (tSSS) and head movement compensation transformed to the mean head position were used first (Neuromag Maxfilter 2.2) to suppress noise from outside of the MEG dewar and effects related to infants' head movement during the recording while minimizing reconstruction error (Taulu & Hari, 2009; Taulu & Kajola, 2005). Then, the signal-space projection method (SSP) was adopted to isolate components of physiological artifacts (i.e. heartbeats and eye blinks), using in-house MATLAB scripts (Uusitalo & Ilmoniemi, 1997). Lastly, the signal was band-pass filtered from 1 to 40Hz and noisy and dead channels were rejected based on the overall power calculated of each channel.

3. Individual analysis

Epoch average

Epochs were rejected when the peak-to-peak amplitude was over 1.5 pT/cm for gradiometers or 2.0 pT/cm for magnetometers. Epochs (-50 to 900ms) in response to standard trials and deviant trials were then averaged separately for each subject after baseline-correction. Baseline-correction was accomplished by subtracting the mean value of the time period prior to trial onset (-50 to 0 ms) from the epoch.

Source Modeling.

Forward modeling employed the Boundary Element Method (BEM) isolated-skull approach with inner skull surface extracted from the MRI of the template. Both the source space and the BEM surface were then aligned and scaled to optimally fit each subject's head shape revealed by head digitization points. All modeling was done with in-house MATLAB scripts in combination with the MNE software (Gramfort et al., 2013).

Inverse source modeling was performed using the dynamic statistic parametric

mapping method (dSPM) without dipole orientation constraints and with data from both gradiometers and magnetometers (Dale et al., 2000). The source activities were normalized to the noise covariance computed from the corresponding empty-room recording, which underwent the same preprocessing steps except for the movement compensation. This procedure resulted in statistically normalized scores for 3 dipole components at each source location for each time point (i.e. dipole strengths in 3 orthogonal directions). The difference between standards and deviants was then computed for each source location at each time point through: 1) subtraction in each of the dipole components (equation 1), and 2) calculating the magnitude of the difference wave (equation 2).

$$[diffx, diffy, diffz] = [devx - standx, devy - standy, devz - stand] \text{ --- (1)}$$

$$maganitude\ of\ difference = \sqrt{(diffx)^2 + (diffy)^2 + (diffz)^2} \text{ --- (2)}$$

4. Group comparison

Each subject's difference waves at all the source locations were interpolated onto a spherical atlas for group level inferences. The FreeSurfer Destrieux atlas was also projected onto this spherical atlas for labeling each source point. Based on the labeling, difference waves in the temporal regions and prefrontal regions were then averaged separately for each subject. The prefrontal regions included superior, middle and inferior areas of the frontal lobe, and the temporal regions included the superior and middle areas of temporal lobes (Figure 4).

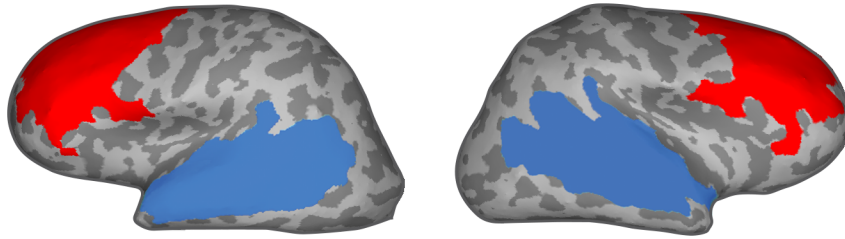


Figure 4 Inflated surface of the head template and the selection for temporal regions (blue) and prefrontal regions (red) for characterizing MMR in these regions.

The Mismatch Responses (MMRs) for temporal and prefrontal regions were calculated for each subject by averaging the response within a target window post violation to best characterize the peak. For the music condition, the target window was 150-300ms post-violation for temporal regions (Figure 5a), and 200-350ms post-violation for the prefrontal regions (Figure 5b). For the speech condition, because the onset of violation (i.e., onset of the second syllable) is 210ms later than the music condition, the target window shifted accordingly in relation to the onset of the trial for both temporal (360-510ms) and prefrontal regions (410-560ms) (Figure 6a, b).

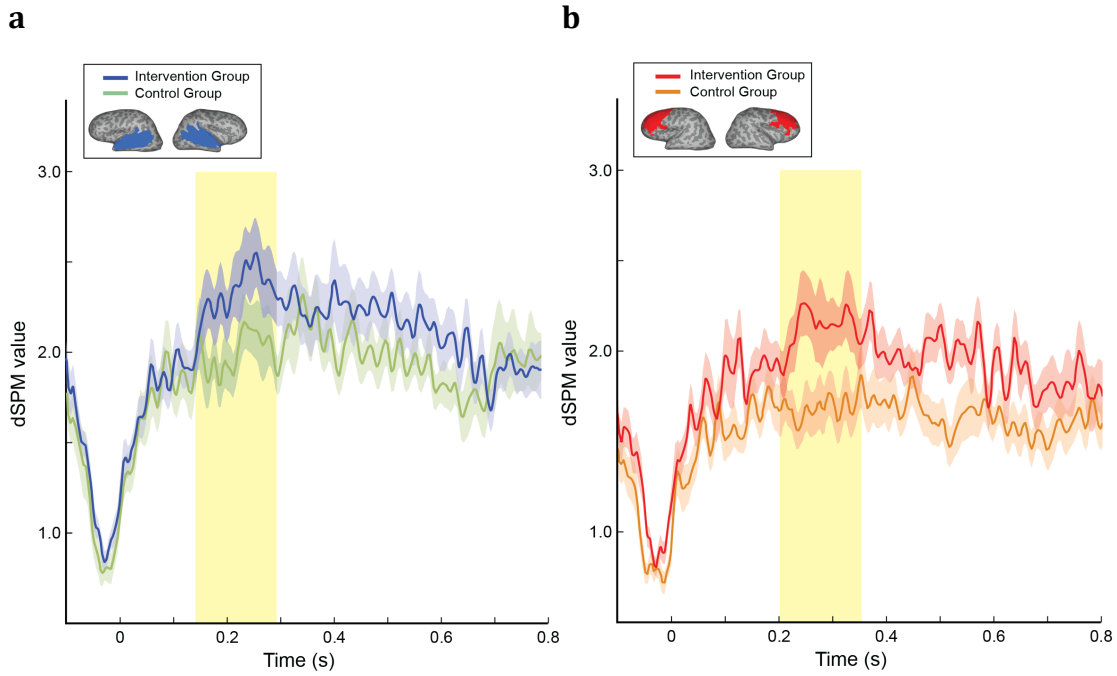


Figure 5 a) Mismatch response (MMR) from the temporal regions for the Intervention group (blue) and the Control group (green) in the music condition; b) Mismatch response (MMR) from the prefrontal regions for the Intervention group (red) and the Control group (orange) in the music condition.

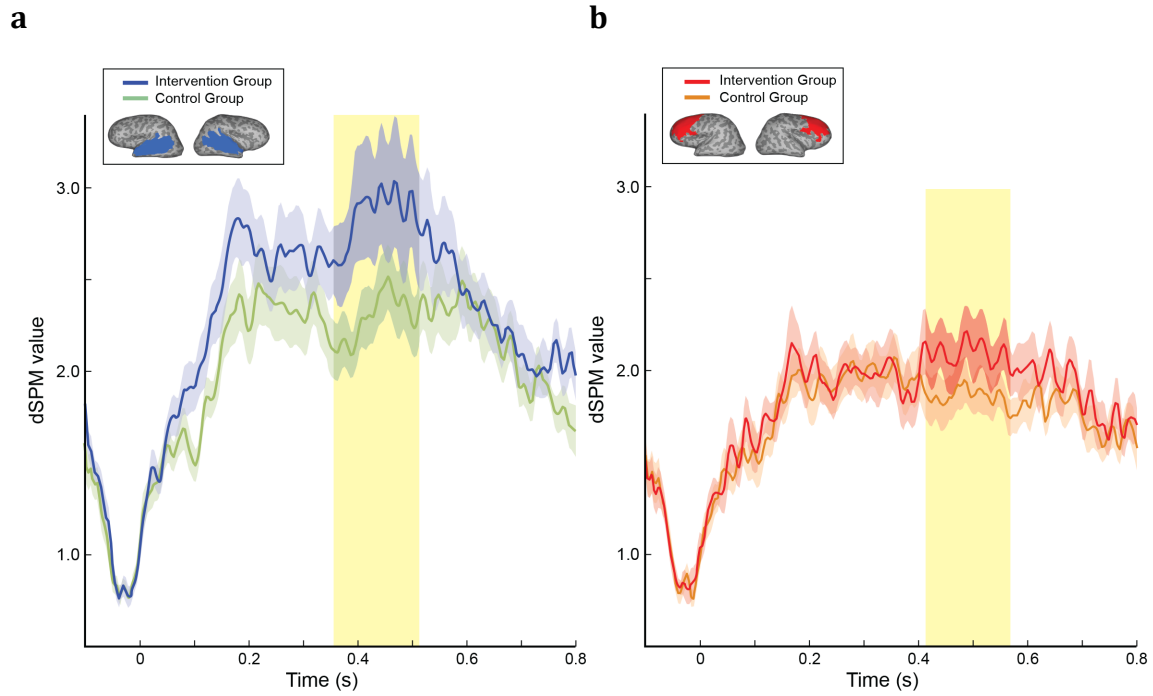


Figure 6 a) Mismatch response from the temporal regions for the Intervention group (blue) and the Control group (green) in the speech condition; b) Mismatch response from the prefrontal regions for the Intervention group (red) and the Control group (orange) in the speech condition.

Chapter III: Results

I. Music Condition

The MMRs were submitted to a mixed 2 (Intervention vs. Control, between subjects) by 2 (Temporal regions vs. Frontal regions, within subjects) ANOVA for statistical comparison (IBM SPSS Inc. Version 19.0). The ANOVA revealed significant main effects of Group, $F(1,34) = 6.29, p=0.017, \eta^2=0.16$, as well as of Region, $F(1,34)=7.32, p=0.011, \eta^2=0.18$. The Intervention group (Mean=2.232, SD=0.11) exhibited larger MMR than the Control group (Mean=1.842, SD=0.11). The temporal regions (Mean=2.169, SD=0.099) exhibited larger MMR than the frontal regions (Mean=1.906, SD=0.084) (Figure 7). No interaction between Group and Region was observed.

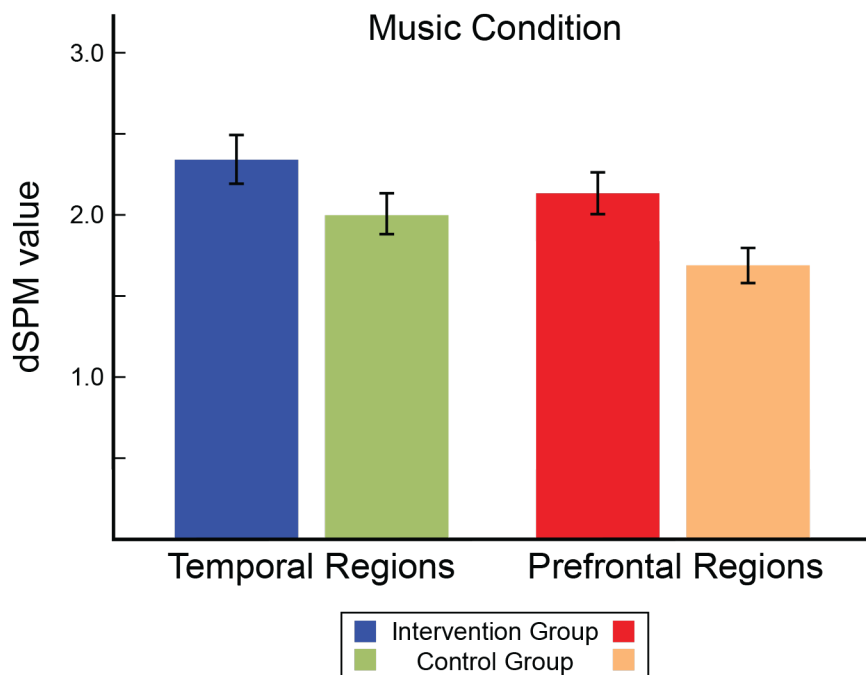


Figure 7 Statistical comparison of MMR in the music condition.

II. Speech Condition

The identical mixed 2 X 2 ANOVA was used to statistically compare the MMRs in the speech condition. As in the music condition, the ANOVA revealed main effects of Group, $F(1,33)=4.56, p=0.039, \eta^2=0.12$, and of Region, $F(1,33)=13.33, p=0.001, \eta^2=0.29$. The Intervention group (Mean=2.42, SD=0.14) exhibited larger MMRs than the Control group (Mean=2.02, SD=0.13). The temporal regions (Mean=2.44, SD=0.14) exhibited larger MMRs than the frontal regions (Mean=1.99, SD=0.08) (Figure 8). No interaction between Groups and Regions was observed.

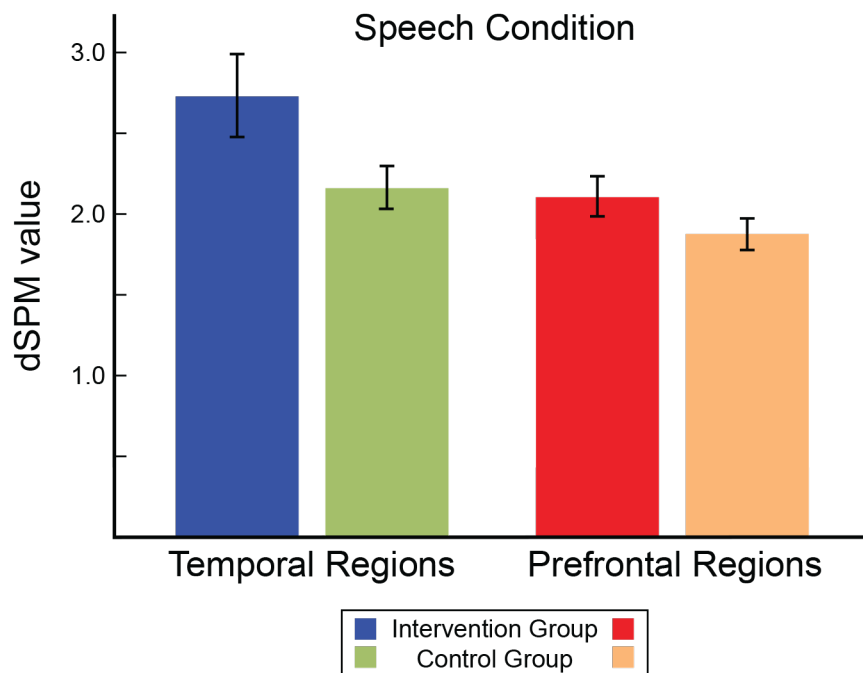


Figure 8 Statistical comparison of MMR in the speech condition.

Chapter IV: Discussion

I. Summary

Little is known about the effects of systematic music intervention on infants in the domain of temporal information processing. The current study aimed to investigate two main questions: 1) the effect of music intervention at 9 month of age on infants' processing of temporal structure in music, and 2) the generalization of the effects to infants' processing of temporal structure in speech. Nine-month-old infants were randomly assigned to complete 12 sessions of structured music intervention targeting triple meter music or 12 control sessions of free play. Mismatch responses (MMRs) were used to quantify their abilities to process the temporal structures in both music and speech following intervention, by measuring sensitivities to temporal structure violation in music and in speech. Previous studies have provided strong evidence for bilateral auditory cortices (temporal regions) as contributing sources to MMR, while evidence of contribution from prefrontal regions was less strong. Therefore, the good spatial resolution of MEG technology allowed the current study to address a secondary question: whether intervention related effects are observed separately in both temporal and prefrontal regions.

Results from the current study provided strong evidence that the short-term music intervention at 9 months of age enhanced infants' ability to process temporal structure in the music domain; and of equal importance, the enhancement generalized to speech processing. In addition, these effects were observed in both temporal and prefrontal regions.

The results complement and extend the literature in two important ways. First, the results validate findings from cross-sectional studies focusing on temporal information processing, demonstrating the difference between musically trained and untrained adults in temporal structure processing in music and the generalization to syllable structure processing in speech. The findings also parallel those from the frequency-processing domain indicating that early music learning enhances information processing embedded in high-level musical structure and generalizes to the corresponding domain of information processing in speech sounds. This similar pattern of results may indicate commonality between frequency and temporal information processing in the central auditory system at higher levels. More interestingly, the generalization effect to the speech domain is robust as it was observed in the same time windows and in the same cortical regions. It is also the first known study reporting a generalization effect from music intervention at 9 months of age. The results therefore generate theoretically and practically interesting questions that will be laid out in detail in the following sections.

‘future-oriented attending,’ to track events occurring at different rates. The ‘analytical’ mode tracks events at a faster rate while the ‘future-oriented’ mode tracks events at a slower rate (i.e., higher level structure). Through maturation and music experience, individuals become more capable of using the ‘future-oriented’ mode to track events at a slower rate, and/or at a higher structural level (Drake, Jones, & Baruch, 2000).

Predictions about future events are manifested through this dynamic attending mechanism by allocating more attentional resources to the important time windows. In fact, various mechanisms have been proposed to describe temporal prediction effects in the sensory cortex on a physiological level. For example, the excitability of sensory neurons is modulated through phase reset and/or interplay between different rates of oscillatory activities (Arnal & Giraud, 2012; Schroeder & Lakatos, 2009). Tuning attention in time enhances the processing of stimuli in the sensory cortex and when the prediction is mismatched with the incoming stimulus, a mismatch response is observed (Rohenkohl, Cravo, Wyart, & Nobre, 2012; Winkler, 2007).

In the current study, repetition of the musical activities as well as multisensory input during music intervention strengthens the dynamic attending mechanism in the Intervention group, allowing faster extraction of temporal structure and stronger predictions in both music and speech. When a prediction was violated in our testing paradigm, a large MMR was therefore observed in response to the violation in the Intervention group, as predicted. In addition, the enhancement of MMR in the prefrontal regions supports the interpretation of a strengthened attentional mechanism, as prefrontal regions have been demonstrated to be involved in attentional networks (Fritz, Elhilali, David, & Shamma, 2007).

There is evidence indicating that enhanced information processing is due to strengthened central auditory processes, supporting our contention that music intervention at 9 months of age enhances these central auditory processes. First, cross-sectional data both from frequency and temporal information processing domains has shown that early music learning related effects are most often observed when target information is embedded within a structure. For example, the ability to detect pitch differences is enhanced in musicians compared to nonmusicians only when within the structure of a chord or a melody, not in pure tone conditions (Koelsch et al., 1999); and similarly, the enhancement in musicians in terms of temporal information processing is observed in metrical structure processing, not note-to-note rhythm (Vuust et al., 2005). There is also evidence arguing directly for enhanced cognitive functions related to early music training, in areas such as attention, working memory and executive function (George & Coch, 2011; Moreno et al., 2011; Moreno & Bidelman, 2014; Sauseng, Griesmayr, Freunberger, & Klimesch, 2010; Schlaug, Norton, Overy, & Winner, 2005; Strait, Kraus, Parbery-Clark, & Ashley, 2010). However, the measurements in these previous studies are generally standardized behavioral measures, which do not allow us to examine a specific attentional mechanism nor its physiological bases. Finally, while the primary system for processing frequency and temporal information matures by the age of 6 months (Olsho, Koch, & Halpin, 1987; Trainor, Samuel, Desjardins, & Sonnadara, 2001; Werner, 2002), the central auditory system appears to undergo a long maturation process extending into young adulthood, and it is this maturation process that is the main contributor to developmental differences from infancy through adulthood,

particularly in challenging environments (e.g. speech perception in noise) (Fallon, Trehub, & Schneider, 2000; Ponton, Eggermont, Kwong, & Don, 2000).

Systematic investigation in the future is warranted to examine this theoretical framework through rigorous testing. The framework provides the means to tap into questions that will allow a deeper understanding of the neural mechanisms shared by speech and music processing and the extent to which they share such mechanisms. Specifically, as a starting point, it is important to investigate how tempo interacts with infants' ability to extract and track temporal structure. The slower the tempo, the more working memory is required to hold information for the temporal structure to be extracted. In fact, Russeler and colleagues (2001) have demonstrated that adult musicians could track isochronous beats at a much slower tempo than the non-trained individuals. Comparable paradigms in music and in speech will be designed in which the tempo (or rate of speech) varies. Infants' ability to track the temporal structure can be measured behaviorally in an adaptive manner as tempo decreases. This method will allow the measurement of the threshold in tempo at which temporal structure can be successfully extracted as well as the modeling of attention from the data. With these measures, several questions can be addressed: 1) What is the developmental trajectory from 6 to 12 months of age in processing temporal structure in music and in speech, as the tempo varies? 2) How does the role of attention change in processing temporal structure in this development window? 3) How does music intervention influence temporal structure processing threshold and attention and how does it interact with development?

Further, the neural correlates allowing for temporal structure tracking can be characterized by measuring infants' neural activities when presented with music or

speech stream above and below the threshold at which they have shown to be able to track the temporal structural information behaviorally. Several candidate electrophysiological measures can be explored to characterize the mechanism: 1) the power at the frequency where temporal structure repeats itself. For example, if a strong beat happens every 500ms, a peak at 2Hz in the oscillatory activities of the brain will be observed. The magnitude of this peak may indicate how well the brain is tracking the regularity (Nozaradan, Peretz, Missal, & Mouraux, 2011). 2) The magnitude of the evoked response and the phase information immediately before the stimulus onset. Previous research has suggested that stimuli that have been predicted to occur at a specific time are processed better than unpredicted ones, manifested by a larger evoked response and phase reset prior to the onset of the stimuli (Arnal & Giraud, 2012; Lange, 2009).

III. Applications of Musical Intervention

Results from the current study also indicate opportunities to examine the effects of music intervention in infants more comprehensively (e.g. long term effects on language development), which may ultimately result in applications as an early clinical intervention method for communication disorders.

Infants begin to learn the speech sounds in their language environments well before the age of 1 year. In particular, the time period between 6 and 12 months of age has been a focus of research, since infants' abilities to process speech sounds diverge depending on whether or not the sounds belong to their native language. Infants' ability to detect changes in nonnative speech sounds declines during this time period, while discrimination ability improves for native speech sounds (Kuhl et al., 2006; Werker & Tees, 1984). In addition, cognitive effort may be reduced in processing native speech sounds in comparison to nonnative speech sounds, as reflected by the cortical oscillatory response in the theta band range (4-8Hz) (Bosseler et al., 2013).

Infants' ability to process speech sounds during this period is critically associated with language acquisition in subsequent development. Behavioral and electrophysiological measures of typically developing infants between 6 and 12 months of age has demonstrated that infants' ability to process acoustic differences in speech sounds predicts the growth of language abilities and language outcomes up to the age of 30 months, particularly vocabulary size (Kuhl et al., 2008; Rivera-Gaxiola, et al., 2005a; Tsao, Liu, & Kuhl, 2004).

Previous research suggested that experience with language between 6-12 months of age influenced infants' speech sound learning (Conboy & Kuhl, 2011; Kuhl et al., 2008;

Kuhl et al., 2003; Roseberry & Kuhl, 2013). Results from the current study suggest that another type of acoustic experience (i.e., music intervention) also influences speech processing during this period. Therefore, it is important to examine whether music intervention in infancy has any long-term effects on language development, as well as investigating any interactions between the effects of language experience and music intervention.

Addressing these questions may shed light on the potential of music intervention as a clinical tool for infants at risk for communication disorders such as dyslexia. Recent studies have argued for an auditory processing deficit component in dyslexia (Loui, Kroog, Zuk, Winner, & Schlaug, 2011; Tallal & Gaab, 2006). More specifically, one theoretical model has argued strongly that deficits in temporal information processing contribute to dyslexia (Goswami, 2011); that is, individuals with dyslexia are impaired in their ability to process the amplitude modulation information in speech, particularly at the slower modulation rates. Thus, music intervention may be useful as an early intervention to mitigate potential deficits in auditory processing in infants at risk for communication disorders such as dyslexia.

A series of specific questions are being investigated as a starting point to gain more comprehensive understanding of music intervention in infancy for potential future applications. 1) Does temporal structure processing (reflected by MMR) at 9-10 month of age predict later language development? 2) Are there any long-term effects of the music intervention on language development at the group level and if so, 3) are the components of the music intervention differentially effective? To characterize language development, we are currently tracking the communication and language skills of both groups of

infants using the MacArthur-Bates Communicative Development Inventory (Fenson et al., 1993), including the non-verbal gestures and size of vocabulary at various ages. To address the first two questions, we will examine differences between the Intervention and Control group in their communication and language skills at different developmental stages as well as the predictive power of the MMR measures in terms of language development.

We hypothesize two specific components to contribute to the effectiveness of music intervention: the amount of synchronized movement and the level of interaction among infants of the sessions. To characterize the amount of synchronized movement to the musical beats, video data from all the music intervention sessions are being analyzed for each infant. This is a particularly interesting component as recent literature has suggested critical involvement of the motor system in keeping timing information, revealed by beta-band oscillatory activities (Fujioka, Trainor, Large, & Ross, 2012). Social factors have also been identified as an important aspect of language learning by previous research (Kuhl et al., 2003) and it can be characterized by measures such as an index that takes into consideration of the number of infants participating in each session. The goal is to evaluate the predictive value of these two variables on both the MMR outcome as well as the long-term language ability measured by CDI using regression models.

IV. Conclusions

The current study demonstrated that short-term music intervention at 9 months of age enhanced infants' neural processing of temporal structure in music, and also that the effect generalized to temporal structure processing in speech. The effects were observed in both temporal and prefrontal regions. A theoretical framework was provided to interpret the effect as a result of an enhanced dynamic attentional system shared by music and speech processing. Future research is warranted to systematically examine this framework to enrich our understanding of shared neural mechanisms between music and speech processing and the plasticity effects related to early intervention. In addition, a more comprehensive understanding of music intervention, such as its long-term effects on language development, is warranted to evaluate the potential for music intervention as an early intervention method for communication disorders.

Reference

- Alho, K. (1995). Cerebral generators of mismatch negativity (MMN) and its magnetic counterpart (MMNM) elicited by sound changes. *Ear and Hearing, 16*(1), 38-51. doi: 10.1097/00003446-199502000-00004
- Arnal, L. H., & Giraud, A. L. (2012). Cortical oscillations and sensory predictions. *Trends in Cognitive Sciences, 16*(7), 390-398. doi: 10.1016/j.tics.2012.05.003
- Basirat, A., Dehaene, S., & Dehaene-Lambertz, G. (2014). A hierarchy of cortical responses to sequence violations in three-month-old infants. *Cognition, 132*(2), 137-150. doi: <http://dx.doi.org/10.1016/j.cognition.2014.03.013>
- Bekinschtein, T. A., Dehaene, S., Rohaut, B., Tadel, F. O., Cohen, L., & Naccache, L. (2009). Neural signature of the conscious processing of auditory regularities. *Proceedings of the National Academy of Sciences of the United States of America, 106*(5), 1672-1677. doi: 10.1073/pnas.0809667106
- Bendixen, A., SanMiguel, I., & Schroger, E. (2012). Early electrophysiological indicators for predictive processing in audition: A review. *International Journal of Psychophysiology, 83*(2), 120-131. doi: 10.1016/j.ijpsycho.2011.08.003
- Bergeson, T. R., & Trehub, S. E. (2006). Infants' perception of rhythmic patterns. *Music Perception, 23*(4), 345-360. doi: 10.1525/mp.2006.23.4.345
- Besson, M., Chobert, J., & Marie, C. I. (2011). Transfer of training between music and speech: common processing, attention and memory. *Frontiers in Psychology, 2*. doi: 10.3389/fpsyg.2011.00094

- Bigand, E., & Poulin-Charronnat, B. (2006). Are we "experienced listeners"? A review of the musical capacities that do not depend on formal musical training. *Cognition, 100*(1), 100-130. doi: 10.1016/j.cognition.2005.11.007
- Boersma, P., & Weenink, D. (2009). Praat: doing phonetics by computer (Version 5.1.05) [Computer program].
- Bosseler, A. N., Taulu, S., Pihko, E., Mäkelä, J. P., Imada, T., Ahonen, A., & Kuhl, P. K. (2013). Theta brain rhythms index perceptual narrowing in infant speech perception. *Front Psychol, 4*. doi: 10.3389/fpsyg.2013.00690
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10*, 433-436.
- Brown, D. E. (1991). *Human universals*. Philadelphia: Temple University Press.
- Chandrasekaran, B., Krishnan, A., & Gandour, J. T. (2009). Relative influence of musical and linguistic experience on early cortical processing of pitch contours. *Brain and Language, 108*(1), 1-9. doi: 10.1016/j.bandl.2008.02.001
- Cheour, M., Alho, K., Ceponiene, R., Reinikainen, K., Sainio, K., Pohjavuori, M., . . . Naatanen, R. (1998). Maturation of mismatch negativity in infants. *International Journal of Psychophysiology, 29*(2), 217-226. doi: 10.1016/s0167-8760(98)00017-8
- Cheour, M., Imada, T., Taulu, S., Ahonen, A., Salonen, J., & Kuhl, P. (2004). Magnetoencephalography is feasible for infant assessment of auditory discrimination. *Experimental Neurology, 190*, S44-S51. doi: 10.1016/j.expneurol.2004.06.030
- Conboy, B. T., & Kuhl, P. K. (2011). Impact of second-language experience in infancy: brain measures of first- and second-language speech perception.

- Developmental Science*, 14(2), 242-248. doi: 10.1111/j.1467-7687.2010.00973.x
- Dale, A. M., Liu, A. K., Fischl, B. R., Buckner, R. L., Belliveau, J. W., Lewine, J. D., & Halgren, E. (2000). Dynamic statistical parametric mapping: combining fMRI and MEG for high-resolution imaging of cortical activity. *Neuron*, 26(1), 55-67.
- Drake, C., Jones, M. R., & Baruch, C. (2000). The development of rhythmic attending in auditory sequences: attunement, referent period, focal attending. *Cognition*, 77(3), 251-288. doi: 10.1016/s0010-0277(00)00106-2
- Essens, P. J. (1986). Hierarchical organization of temporal patterns. *Perception & Psychophysics*, 40(2), 69-73. doi: 10.3758/bf03208185
- Fallon, M., Trehub, S. E., & Schneider, B. A. (2000). Children's perception of speech in multitalker babble. *Journal of the Acoustical Society of America*, 108(6), 3023-3029. doi: 10.1121/1.1323233
- Fenson, L., Dale, P., Reznick, J. S., Thal, D., Bates, E., Hartung, J., . . . Reilly, J. S. (1993). *MacArthur communicative development inventories: user's guide and technical manual*: San Diego, CA: Singular Publishing Group.
- Fraisse, P. (1982). Rhythm and tempo. In D. Deutsch (Ed.), *The Psychology of Music*. New York: Academy Press.
- Fritz, J. B., Elhilali, M., David, S. V., & Shamma, S. A. (2007). Auditory attention - focusing the searchlight on sound. *Current Opinion in Neurobiology*, 17(4), 437-455. doi: 10.1016/j.conb.2007.07.011

- Fujioka, T., Ross, B., Kakigi, R., Pantev, C., & Trainor, L. J. (2006). One year of musical training affects development of auditory cortical-evoked fields in young children. *Brain*, *129*, 2593-2608. doi: 10.1093/brain/aw1247
- Fujioka, T., Trainor, L. J., Large, E. W., & Ross, B. (2012). Internalized Timing of Isochronous Sounds Is Represented in Neuromagnetic Beta Oscillations. *Journal of Neuroscience*, *32*(5), 1791-1802. doi: 10.1523/jneurosci.4107-11.2012
- Fujioka, T., Trainor, L. J., Ross, B., Kakigi, R., & Pantev, C. (2004). Musical training enhances automatic encoding of melodic contour and interval structure. *Journal of Cognitive Neuroscience*, *16*(6), 1010-1021. doi: 10.1162/0898929041502706
- Fujioka, T., Trainor, L. J., Ross, B., Kakigi, R., & Pantev, C. (2005). Automatic encoding of polyphonic melodies in musicians and nonmusicians. *Journal of Cognitive Neuroscience*, *17*(10), 1578-1592. doi: 10.1162/089892905774597263
- Garcia-Sierra, A., Rivera-Gaxiola, M., Percaccio, C. R., Conboy, B. T., Romo, H., Klarman, L., . . . Kuhl, P. K. (2011). Bilingual language learning: An ERP study relating early brain responses to speech, language input, and later word production. *Journal of Phonetics*, *39*(4), 546-557. doi: 10.1016/j.wocn.2011.07.002
- Geiser, E., Sandmann, P., Jancke, L., & Meyer, M. (2010). Refinement of metre perception - training increases hierarchical metre processing. *European Journal of Neuroscience*, *32*(11), 1979-1985. doi: 10.1111/j.1460-9568.2010.07462.x

- George, E. M., & Coch, D. (2011). Music training and working memory: An ERP study. *Neuropsychologia*, *49*(5), 1083-1094. doi: 10.1016/j.neuropsychologia.2011.02.001
- 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. doi: 10.1111/j.1467-7687.2012.01142.x
- Giard, M. H., Perrin, F., Pernier, J., & Bouchet, P. (1990). Brain generators implicated in the processing of auditory stimulus deviance: a topographic event-related potential study. *Psychophysiology*, *27*(6), 627-640.
- Gomot, M., Giard, M. H., Roux, S., Barthelemy, C., & Bruneau, N. (2000). Maturation of frontal and temporal components of mismatch negativity (MMN) in children. *Neuroreport*, *11*(14), 3109-3112. doi: 10.1097/00001756-200009280-00014
- Goswami, U. (2011). A temporal sampling framework for developmental dyslexia. *Trends in Cognitive Sciences*, *15*(1), 3-10. doi: 10.1016/j.tics.2010.10.001
- Goswami, U., & Leong, V. (2013). Speech rhythm and temporal structure: Converging perspectives? *Laboratory Phonology*, *4*(1), 67-92. doi: 10.1515/lp-2013-0004
- Gramfort, A., Luessi, M., Larson, E., Engemann, D., Strohmeier, D., Brodbeck, C., . . . Hamalainen, M. (2013). MNE software for processing MEG and EEG data. *Neuroimage*, *86*, 446-460.
- Hamalainen, M., Hari, R., Ilmoniemi, R. J., Knuutila, J., & Lounasmaa, O. V. (1993). Magnetoencephalography, theory, instrumentation, and applications to noninvasive studies of the working human brain. *Reviews of Modern Physics*, *65*(2), 413-497.

Herholz, Sibylle C., & Zatorre, Robert J. (2012). Musical Training as a Framework for Brain Plasticity: Behavior, Function, and Structure. *Neuron*, 76(3), 486-502.

Honing, H. (2012). Without it no music: beat induction as a fundamental musical trait *Neurosciences and Music Iv: Learning and Memory* (Vol. 1252, pp. 85-91). Oxford: Blackwell Science Publ.

Imada, T., Zhang, Y., Cheour, M., Taulu, S., Ahonen, A., & Kuhl, P. K. (2006). Infant speech perception activates Broca's area: a developmental magnetoencephalography study. *Neuroreport*, 17(10), 957-962. doi: 10.1097/01.wnr.0000223387.51704.89

Jones, M. R., & Boltz, M. (1989). DYNAMIC ATTENDING AND RESPONSES TO TIME. *Psychological Review*, 96(3), 459-491. doi: 10.1037//0033-295x.96.3.459

Koelsch, S., Schroger, E., & Tervaniemi, M. (1999). Superior pre-attentive auditory processing in musicians. *Neuroreport*, 10(6), 1309-1313.

Kraus, N., & Chandrasekaran, B. (2010). Music training for the development of auditory skills. *Nature Reviews Neuroscience*, 11(8), 599-605. doi: 10.1038/nrn2882

Kuhl, P. K. (1998). *Effects of language experience on speech perception*. Paper presented at the Acoustical Society of America, Woodbury, NY.

Kuhl, P. K., Conboy, B. T., Coffey-Corina, S., Padden, D., Rivera-Gaxiola, M., & Nelson, T. (2008). Phonetic learning as a pathway to language: new data and native language magnet theory expanded (NLM-e). *Philosophical Transactions of the Royal Society B-Biological Sciences*, 363(1493), 979-1000. doi: 10.1098/rstb.2007.2154

- Kuhl, P. K., Ramirez, R. R., Bosseler, A., Lin, J.-F. L., & Imada, T. (2014). Infants' brain responses to speech suggest Analysis by Synthesis. *Proceedings of the National Academy of Sciences*. doi: 10.1073/pnas.1410963111
- Kuhl, P. K., Stevens, E., Hayachi, A., Deguchi, T., Kiritani, S., & Iverson, P. (2006). Infants show a facilitation effect for native language phonetic perception between 6 and 12 months. *Developmental Science*, 9, F13-F21. doi: doi:10.1111/j.1467-7687.2006.00468.x
- Kuhl, P. K., Tsao, F. M., & Liu, H. M. (2003). Foreign-language experience in infancy: Effects of short-term exposure and social interaction on phonetic learning. *Proceedings of the National Academy of Sciences of the United States of America*, 100(15), 9096-9101. doi: 10.1073/pnas.1532872100
- Lange, K. (2009). Brain correlates of early auditory processing are attenuated by expectations for time and pitch. *Brain and Cognition*, 69(1), 127-137. doi: 10.1016/j.bandc.2008.06.004
- Lappe, C., Herholz, S. C., Trainor, L. J., & Pantev, C. (2008). Cortical plasticity induced by short-term unimodal and multimodal musical training. *Journal of Neuroscience*, 28(39), 9632-9639. doi: 10.1523/jneurosci.2254-08.2008
- Lappe, C., Trainor, L. J., Herholz, S. C., & Pantev, C. (2011). Cortical Plasticity Induced by Short-Term Multimodal Musical Rhythm Training. *Plos One*, 6(6). doi: e2149310.1371/journal.pone.0021493
- Loui, P., Kroog, K., Zuk, J., Winner, E., & Schlaug, G. (2011). Relating Pitch Awareness to Phonemic Awareness in Children: Implications for Tone-Deafness and Dyslexia. *Frontiers in Psychology*, 2. doi: 10.3389/fpsyg.2011.00111

- Magne, C., Schon, D., & Besson, M. (2006). Musician children detect pitch violations in both music and language better than nonmusician children: Behavioral and electrophysiological approaches. *Journal of Cognitive Neuroscience, 18*(2), 199-211.
- Marie, C., Delogu, F., Lampis, G., Belardinelli, M. O., & Besson, M. (2011). Influence of Musical Expertise on Segmental and Tonal Processing in Mandarin Chinese. *Journal of Cognitive Neuroscience, 23*(10), 2701-2715. doi: 10.1162/jocn.2010.21585
- Marie, C., Magne, C., & Besson, M. (2011). Musicians and the Metric Structure of Words. *Journal of Cognitive Neuroscience, 23*(2), 294-305. doi: 10.1162/jocn.2010.21413
- Martin, J. G. (1972). RHYTHMIC (HIERARCHICAL) VERSUS SERIAL STRUCTURE IN SPEECH AND OTHER BEHAVIOR. *Psychological Review, 79*(6), 487-&. doi: 10.1037/h0033467
- McMullen, E., & Saffran, J. R. (2004). Music and language: A developmental comparison. *Music Perception, 21*(3), 289-311. doi: 10.1525/mp.2004.21.3.289
- Moreno, S., & Besson, M. (2006). Musical training and language-related brain electrical activity in children. *Psychophysiology, 43*(3), 287-291. doi: 10.1111/j.1469-8986.2006.00401.x
- Moreno, S., Bialystok, E., Barac, R., Schellenberg, E. G., Cepeda, N. J., & Chau, T. (2011). Short-Term Music Training Enhances Verbal Intelligence and Executive

- Function. *Psychological Science*, 22(11), 1425-1433. doi:
10.1177/0956797611416999
- Moreno, S., & Bidelman, G. M. (2014). Examining neural plasticity and cognitive benefit through the unique lens of musical training. *Hearing Research*, 308(0), 84-97. doi: <http://dx.doi.org/10.1016/j.heares.2013.09.012>
- Munte, T. F., Altenmuller, E., & Jancke, L. (2002). The musician's brain as a model of neuroplasticity. *Nature Reviews Neuroscience*, 3(6), 473-478. doi:
10.1038/nrn843
- Musacchia, G., Sams, M., Skoe, E., & Kraus, N. (2007). Musicians have enhanced subcortical auditory and audiovisual processing of speech and music. *Proceedings of the National Academy of Sciences of the United States of America*, 104(40), 15894-15898. doi: 10.1073/pnas.0701498104
- Musacchia, G., Strait, D., & Kraus, N. (2008). Relationships between behavior, brainstem and cortical encoding of seen and heard speech in musicians and non-musicians. *Hearing Research*, 241(1-2), 34-42. doi:
10.1016/j.heares.2008.04.013
- Naatanen, R., Paavilainen, P., Rinne, T., & Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: A review. *Clinical Neurophysiology*, 118(12), 2544-2590. doi: 10.1016/j.clinph.2007.04.026
- Nozaradan, S., Peretz, I., Missal, M., & Mouraux, A. (2011). Tagging the Neuronal Entrainment to Beat and Meter. *Journal of Neuroscience*, 31(28), 10234-10240. doi: 10.1523/jneurosci.0411-11.2011

- Olsho, L. W., Koch, E. G., & Halpin, C. F. (1987). LEVEL AND AGE EFFECTS IN INFANT FREQUENCY DISCRIMINATION. *Journal of the Acoustical Society of America*, *82*(2), 454-464. doi: 10.1121/1.395446
- Ortiz-Mantilla, S., Hamalainen, J. A., & Benasich, A. A. (2012). Time course of ERP generators to syllables in infants: A source localization study using age-appropriate brain templates. *Neuroimage*, *59*(4), 3275-3287. doi: 10.1016/j.neuroimage.2011.11.048
- Pantev, C., Oostenveld, R., Engelien, A., Ross, B., Roberts, L. E., & Hoke, M. (1998). Increased auditory cortical representation in musicians. *Nature*, *392*(6678), 811-814. doi: 10.1038/33918
- Pantev, C., Roberts, L. E., Schulz, M., Engelien, A., & Ross, B. (2001). Timbre-specific enhancement of auditory cortical representations in musicians. *Neuroreport*, *12*(1), 169-174. doi: 10.1097/00001756-200101220-00041
- Parbery-Clark, A., Skoe, E., Lam, C., & Kraus, N. (2009). Musician Enhancement for Speech-In-Noise. *Ear and Hearing*, *30*(6), 653-661.
- Parbery-Clark, A., Tierney, A., Strait, D. L., & Kraus, N. (2012). Musicians have fine-tuned neural distinctions of speech syllables. *Neuroscience*, *219*, 111-119. doi: 10.1016/j.neuroscience.2012.05.042
- Patel, A. D. (2008). *Music, language, and the brain*. Oxford; New York: Oxford University Press.
- Phillips-Silver, J., & Trainor, L. J. (2005). Feeling the beat: Movement influences infant rhythm perception. *Science*, *308*(5727), 1430-1430. doi: 10.1126/science.1110922

- Phillips-Silver, J., & Trainor, L. J. (2007). Hearing what the body feels: Auditory encoding of rhythmic movement. *Cognition*, *105*(3), 533-546. doi: 10.1016/j.cognition.2006.11.006
- Phillips-Silver, J., & Trainor, L. J. (2008). Vestibular influence on auditory metrical interpretation. *Brain and Cognition*, *67*(1), 94-102. doi: 10.1016/j.bandc.2007.11.007
- Poeppl, D., Idsardi, W. J., & van Wassenhove, V. (2008). Speech perception at the interface of neurobiology and linguistics. *Philosophical Transactions of the Royal Society B-Biological Sciences*, *363*(1493), 1071-1086. doi: 10.1098/rstb.2007.2160
- Ponton, C. W., Eggermont, J. J., Kwong, B., & Don, M. (2000). Maturation of human central auditory system activity: evidence from multi-channel evoked potentials. *Clinical Neurophysiology*, *111*(2), 220-236. doi: 10.1016/s1388-2457(99)00236-9
- Povel, D. J., & Essens, P. (1985). PERCEPTION OF TEMPORAL PATTERNS. *Music Perception*, *2*(4), 411-440.
- Rivera-Gaxiola, M., Klarman, L., Garcia-Sierra, A., & Kuhl, P. K. (2005). Neural patterns to speech and vocabulary growth in American infants. *Neuroreport*, *16*(5), 495-498. doi: 10.1097/00001756-200504040-00015
- Rivera-Gaxiola, M., Silva-Pereyra, J., & Kuhl, P. K. (2005). Brain potentials to native and non-native speech contrasts in 7-and 11-month-old American infants. *Developmental Science*, *8*(2), 162-172. doi: 10.1111/j.1467-7687.2005.00403.x

- Rohenkohl, G., Cravo, A. M., Wyart, V., & Nobre, A. C. (2012). Temporal Expectation Improves the Quality of Sensory Information. *Journal of Neuroscience*, *32*(24), 8424-8428. doi: 10.1523/jneurosci.0804-12.2012
- Rosburg, T., Trautner, P., Dietl, T., Korzyukov, O. A., Boutros, N. N., Schaller, C., . . . Kurthen, M. (2005). Subdural recordings of the mismatch negativity (MMN) in patients with focal epilepsy. *Brain*, *128*, 819-828. doi: 10.1093/brain/awh442
- Roseberry, S., & Kuhl, P. (2013). *Two are better than one: Infant language learning in the presence of peers.* . Paper presented at the The Society for Research in Child Development, Seattle WA.
- Rosen, S. (1992). Temporal information in speech - acoustic, auditory and linguistic aspects. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences*, *336*(1278), 367-373. doi: 10.1098/rstb.1992.0070
- Russeler, J., Altenmüller, E., Nager, W., Kohlmetz, C., & Munte, T. (2001). Event-related brain potentials to sound omissions differ in musicians and non-musicians. *Neuroscience Letters*, *308*(1), 33-36. doi: 10.1016/s0304-3940(01)01977-2
- Sauseng, P., Griesmayr, B., Freunberger, R., & Klimesch, W. (2010). Control mechanisms in working memory: A possible function of EEG theta oscillations. *Neuroscience and Biobehavioral Reviews*, *34*(7), 1015-1022. doi: 10.1016/j.neubiorev.2009.12.006

- Schlaug, G. (2001). The brain of musicians: A model for functional and structural adaptation. In R. J. Zatorre & I. Peretz (Eds.), *Biological Foundations of Music* (Vol. 930, pp. 281-299). New York: New York Acad Sciences.
- Schlaug, G., Jancke, L., Huang, Y. X., Staiger, J. F., & Steinmetz, H. (1995). Increased Corpus-Callosum size in musicians. *Neuropsychologia*, *33*(8), 1047-&. doi: 10.1016/0028-3932(95)00045-5
- Schlaug, G., Norton, A., Overy, K., & Winner, E. (2005). Effects of music training on the child's brain and cognitive development. In G. Avanzini, S. Koelsch, L. Lopez & M. Majno (Eds.), *Neurosciences and Music II: From Perception to Performance* (Vol. 1060, pp. 219-230). New York: New York Acad Sciences.
- 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. doi: 10.1038/nn871
- Schon, D., Magne, C., & Besson, M. (2004). The music of speech: Music training facilitates pitch processing in both music and language. *Psychophysiology*, *41*(3), 341-349. doi: 10.1111/1469-8986.00172.x
- Schroeder, C. E., & Lakatos, P. (2009). Low-frequency neuronal oscillations as instruments of sensory selection. *Trends in Neurosciences*, *32*(1), 9-18. doi: 10.1016/j.tins.2008.09.012
- Shahin, A. J., Roberts, L. E., Chau, W., Trainor, L. J., & Miller, L. M. (2008). Music training leads to the development of timbre-specific gamma band activity. *Neuroimage*, *41*(1), 113-122. doi: 10.1016/j.neuroimage.2008.01.067

- Shannon, R. V., Zeng, F. G., Kamath, V., Wygonski, J., & Ekelid, M. (1995). Speech recognition with primarily temporal cues. *Science*, *270*(5234), 303-304. doi: 10.1126/science.270.5234.303
- Strait, D. L., Kraus, N., Parbery-Clark, A., & Ashley, R. (2010). Musical experience shapes top-down auditory mechanisms: Evidence from masking and auditory attention performance. *Hearing Research*, *261*(1-2), 22-29. doi: 10.1016/j.heares.2009.12.021
- Tallal, P., & Gaab, N. (2006). Dynamic auditory processing, musical experience and language development. *Trends in Neurosciences*, *29*(7), 382-390. doi: 10.1016/j.tins.2006.06.003
- Taulu, S., & Hari, R. (2009). Removal of magnetoencephalographic artifacts with temporal signal-space separation: Demonstration with single-trial auditory-evoked responses. *Human Brain Mapping*, *30*(5), 1524-1534. doi: 10.1002/hbm.20627
- Taulu, S., & Kajola, M. (2005). Presentation of electromagnetic multichannel data: The signal space separation method. *Journal of Applied Physics*, *97*(12). doi: 10.1063/1.1935742
- Tervaniemi, M., Sannemann, C., Noyranen, M., Salonen, J., & Pihko, E. (2011). Importance of the left auditory areas in chord discrimination in music experts as demonstrated by MEG. *European Journal of Neuroscience*, *34*(3), 517-523. doi: 10.1111/j.1460-9568.2011.07765.x

- Trainor, L. J., Samuel, S. S., Desjardins, R. N., & Sonnadara, R. R. (2001). Measuring temporal resolution in infants using mismatch negativity. *Neuroreport*, *12*(11), 2443-2448. doi: 10.1097/00001756-200108080-00031
- Trehub, S. E. (1993). Temporal Auditory Processing in Infancy. *NYAS Annals of the New York Academy of Sciences*, *682*(1), 137-149.
- Trehub, S. E., & Hannon, E. E. (2006). Infant music perception: Domain-general or domain-specific mechanisms? *Cognition*, *100*(1), 73-99. doi: 10.1016/j.cognition.2005.11.006
- Tsao, F. M., Liu, H. M., & Kuhl, P. K. (2004). Speech perception in infancy predicts language development in the second year of life: A longitudinal study. *Child Development*, *75*(4), 1067-1084. doi: 10.1111/j.1467-8624.2004.00726.x
- Uusitalo, M. A., & Ilmoniemi, R. J. (1997). Signal-space projection method for separating MEG or EEG into components. *Medical and Biological Engineering and Computing*, *35*(2), 135-140. doi: 10.1007/BF02534144
- Vuust, P., Pallesen, K. J., Bailey, C., van Zuijen, T. L., Gjedde, A., Roepstorff, A., & Ostergaard, L. (2005). To musicians, the message is in the meter - Pre-attentive neuronal responses to incongruent rhythm are left-lateralized in musicians. *Neuroimage*, *24*(2), 560-564. doi: 10.1016/j.neuroimage.2004.08.039
- Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: evidence for perceptual reorganization during the first year of life. *Infant Behavior & Development*(7), 49-63. doi: doi:10.1016/S0163-6383(84)80022-3

- Werner, L. A. (2002). Infant Auditory Capabilities. *Current Opinion in Otolaryngology & Head and Neck Surgery*, *10*, 398-402.
- Winkler, I. (2007). Interpreting the mismatch negativity. *Journal of Psychophysiology*, *21*(3-4), 147-163. doi: 10.1027/0269-8803.21.34.147
- Winkler, I., Denham, S. L., & Nelken, I. (2009). Modeling the auditory scene: predictive regularity representations and perceptual objects. *Trends in Cognitive Sciences*, *13*(12), 532-540. doi: 10.1016/j.tics.2009.09.003
- Wong, P. C. M., Roy, A. K., & Margulis, E. H. (2009). Bimusicalism: the implicit dual enculturation of cognitive and affective systems. *Music Perception*, *27*(2), 81-88. doi: 10.1525/mp.2009.27.2.81
- Wong, P. C. M., Skoe, E., Russo, N. M., Dees, T., & Kraus, N. (2007). Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nature Neuroscience*, *10*(4), 420-422. doi: 10.1038/nn1872
- Zatorre, R. J. (2013). Predispositions and Plasticity in Music and Speech Learning: Neural Correlates and Implications. *Science*, *342*(6158), 585-589. doi: 10.1126/science.1238414
- Zhang, Y. Y., Brady, M., & Smith, S. (2001). Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. *Ieee Transactions on Medical Imaging*, *20*(1), 45-57. doi: 10.1109/42.906424