

5-2019

Music, Language, and Rhythmic Timing

Rhimmon Simchy-Gross
University of Arkansas, Fayetteville

Follow this and additional works at: <https://scholarworks.uark.edu/etd>

Part of the [Cognition and Perception Commons](#), [Cognitive Psychology Commons](#), and the [Experimental Analysis of Behavior Commons](#)

Recommended Citation

Simchy-Gross, Rhimmon, "Music, Language, and Rhythmic Timing" (2019). *Theses and Dissertations*. 3175.
<https://scholarworks.uark.edu/etd/3175>

This Dissertation is brought to you for free and open access by ScholarWorks@UARK. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of ScholarWorks@UARK. For more information, please contact ccmiddle@uark.edu.

Music, Language, and Rhythmic Timing

A thesis submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy in Psychology

by

Rhimmon Simchy-Gross
The State University of New York at Potsdam
Bachelor of Music in Psychology and Piano Performance, 2013
University of Arkansas
Master of Arts in Psychology, 2016

May 2019
University of Arkansas

This dissertation is approved for recommendation to the Graduate Council.

Elizabeth Hellmuth Margulis, Ph.D.
Dissertation Director

Douglas Behrend, Ph.D.
Committee Member

William Levine, Ph.D.
Committee Member

Abstract

Neural, perceptual, and cognitive oscillations synchronize with rhythmic events in both speech (Luo & Poeppel, 2007) and music (Snyder & Large, 2005). This synchronization decreases perceptual thresholds to temporally predictable events (Lawrance et al., 2014), improves task performance (Ellis & Jones, 2010), and enables speech intelligibility (Peelle & Davis, 2012). Despite implications of music-language transfer effects for improving language outcomes (Gordon et al., 2015), proposals that shared neural and cognitive resources underlie music and speech rhythm perception (e.g., Tierney & Kraus, 2014) are not yet substantiated. The present research aimed to explore this potential overlap by testing whether music-induced oscillations affect metric speech tempo perception, and vice versa. We presented in each of 432 trials a prime sequence (seven repetitions of either a metric speech utterance or analogous musical phrase) followed by a standard-comparison pair (either two identical speech utterances or two identical musical phrases). Twenty-two participants judged whether the comparison was slower than, faster than, or the same tempo as the standard. We manipulated whether the prime was slower than, faster than, or the same tempo as the standard. Tempo discrimination accuracy was higher when the standard tempo was the same as, compared to slower or faster than, the prime tempo. These findings support the shared-resources view more than the independent-resources view, and they have implications for music-language transfer effects showing improvements in verbal memory (Chan et al., 1998), speech-in-noise perception (Strait et al., 2012), and reading ability in children and adults (Tierney & Kraus, 2013).

Acknowledgments

I could not have accomplished all I have accomplished without the unending support of my wife and the endless cuddles of my two miniature schnauzer puppies. I am eternally grateful to my entire family for always believing in me and encouraging me to never stop pursuing my goals.

I could not have been more fortunate to have Dr. Elizabeth Hellmuth Margulis as a mentor over the past six years. She is incredible. I thank her for guiding me through this project at every stage and always bringing out the best in me. She has had an extraordinary influence on me and has been instrumental in my success.

I am very appreciative of Dr. Douglas Behrend and Dr. William Levine for their sustained and meaningful support over the past six years and throughout the development and completion of this project. Their thoughtful feedback drives me to continue looking for ways to make impactful contributions to the literature and push the boundaries in my field.

Table of Contents

Introduction	1
Language, Music, and Rhythm	2
Entrainment in speech	3
Entrainment in music	4
Entrainment in speech and music	6
Shared resources	6
Independent resources.....	9
The Present Research	11
Method	12
Participants	12
Design	12
Stimuli	13
Speech stimulus	14
Music stimulus	15
Procedure	15
Data Analysis	17
Results	17
Discussion	20
Prime tempo asymmetry	21
Other considerations	24
Implications	24
Ecological validity	24
Interval timing and rhythmic timing	25
Music training and language processing	27
Conclusion	28
Figures.....	30
References.....	33
Appendix.....	41

Introduction

Music and language both form complex auditory patterns that have syntactic structure, prosodic variation, and unfold dynamically in time. They both manipulate expectations in powerful ways, and they are both fundamental to communication. Similarities between how humans *perceive* music and language, however, is a matter of current theoretical debate.

In recent years, researchers have focused much of their inquiry into potential music-language overlaps on the question of whether musical syntax is processed by the same cognitive and neural mechanisms as linguistic syntax. Findings on this subject have been contradictory, leading to competing theories. Peretz and Coltheart (2003) suggested that music and language processing are likely domain-specific, have neural specificity, and are governed by separate music-processing and language-processing modules. In support of this view, they highlighted the fact that people can have selective impairment of language processing but not music processing, and vice versa. In contrast, Slevc, Rosenberg and Patel (2009) showed that musical expectancy violations enhanced the disruptive effects of linguistic expectancy violations, suggesting music and speech syntactic processing draw from a limited pool of resources. Patel (2003, 2008) showed that both musical and linguistic syntactic violations produce the same P600 activation, suggesting the existence of significant overlap between the processing of musical and linguistic syntax. Slevc and Okada (2015) argue that musical and linguistic syntactical processing rely not on a dedicated syntax processor but rather on shared mechanisms for cognitive control. Other research has investigated the perceptual overlap of music and language by integrating developmental, clinical, and cross-species perspectives. Here, more evidence for the domain-general view than the domain-specific view emerges (Goswami, 2012). Indeed, relations

between speech and music processing might be due to general perceptual mechanisms that are both domain-general and species-general (Trehub & Hannon, 2006).

One way to help resolve this music-language-overlap debate is by examining the relationship between the perception of musical and linguistic rhythm, instead of syntax.

Language, Music, and Rhythm

Neural oscillations, or cyclic pulses of neural excitability, underlie the perception of temporal input (Buzsáki & Draguhn, 2004). Stimuli whose rhythmic properties occur in phase with oscillatory peaks receive increased attentional focus, decreased perceptual thresholds, and improved task performance (e.g., Jones & Boltz, 1989; Large & Jones, 1999; Lawrance, Harper, Cooke, & Schnupp, 2014). Evidence suggests entrainment does not require consciousness (Oever, 2017), yet can be induced consciously by using predictive cues (Graber & Fujioka, 2017) and imagining different meters (Nozaradan, Peretz, & Mouraux, 2012). Evidence also suggests perceptual benefits can occur even in the absence of external rhythmic stimulation (Henry, Herrmann, & Obleser, 2016). Internal oscillations can be induced experimentally using transcranial alternating current stimulation, fMRI, MEG, and EEG (Härle, Rockstroh, Keil, Wienbruch, & Elbert, 2004; Nozaradan et al., 2012; Zoefel, Archer-Boyd, & Davis, 2017). Furthermore, entrainment effects have been demonstrated in a variety of modalities, domains, and animals (Ammirante, Patel, & Russo, 2016; Bauer, Bleichner, Baillet, & Debener, 2017; Grahn, Henry, & McAuley, 2011; Lakatos, Karmos, Mehta, Ulbert, & Schroeder, 2008; Lakatos, Shah, Knuth, Ulbert, Karmos, & Schroeder, 2005; Luo, Liu, & Poeppel, 2010; McAuley & Henry, 2010; McAuley, Henry, & Tkach, 2012; McAuley, Wang, Miller, & Pang, 2010; Miller, Carlson, & McAuley, 2013; Morillon, Schroeder, & Wyart, 2014; Ng, Logothetis, & Kayser,

2012; Pasinski, McAuley, & Snyder, 2016; Schroeder, Lakatos, Kajikawa, Partan, & Puce, 2008; ten Oever, Schroeder, Poeppel, van Atteveldt, & Zion-Golumbic, 2014).

In the auditory domain, researchers have experimentally induced internal oscillations using external stimuli. Behavioral evidence for entrainment effects have been demonstrated using a priming paradigm (Barnes & Jones, 2000; Large & Jones, 1999; McAuley & Jones, 2003; McAuley & Kidd, 1998). Participants typically hear (although told to ignore) a series of several isochronous or metric prime tones followed by test tones comprising standard and comparison intervals. Participants perform a perceptual discrimination task on the test tones. A duration discrimination task, for example, involves judging whether the ‘comparison’ interval (the duration between the two discrete ‘comparison’ tones) is shorter, longer, or the same duration as the ‘standard’ interval.

Performance is better when the standard interval occurs in phase (versus out of phase) with the prime sequence rhythm—when the ending of the standard occurs on time (versus early or late) relative to the beat of the prime sequence. It is theorized that performance improves because attentional focus and neural activity increase during expected temporal locations (McAuley & Jones, 2003). Neural oscillations (waves of increases and decreases in field potentials or neural firing rates; Ng et al., 2012) synchronize with the onsets of the sine tones. Any stimuli following the prime sequence that occur in phase with the prime rhythm, and thus the entrained neural oscillations, receive perceptual enhancement.

Entrainment in speech. Entrainment is thought to be critical for speech perception (Goswami, 2011; Giraud & Poeppel, 2012; Nazzi, Bertoncini, & Mehler, 1998; Peele & Davis, 2012; Quené & Port, 2005). Synchronization of external, stimulus-induced oscillations and internal, neural oscillations is believed to be a prerequisite for successful language processing

(Kotz, Schwartze, & Schmidt-Kassow, 2009). This synchronization is thought to have a causal influence on the neural processing of intelligible speech (Zoefel, Archer-Boyd, & Davis, 2017). Entrainment is also hypothesized to play an important role in temporal coordination, interpersonal synchrony, and dyadic spoken communication. These interpersonal factors have been shown to improve speech processing and comprehension (Giles, Coupland & Coupland, 1991), enhance social cooperation (Manson, Bryant, Gervais, & Kline, 2013), and facilitate memory, perceptual fluency, and feelings of flow and harmony (Hutchins, 1995; Semin & Cacioppo, 2008).

Although significant evidence suggests entrainment is crucial for successful speech perception, some researchers doubt that temporal prediction is necessary for speech learning, perception, and comprehension (Huetting & Mani, 2016). Similarly, some researchers question whether entrainment underlies the processing of temporally irregular, in addition to regular, speech (Beier & Ferreira, 2018), although people have been shown to entrain to temporally irregular tone sequences (Herrmann, Henry, Haegens, & Obleser, 2016).

Entrainment in music. Entrainment is central to music perception and production, and to the experience of feeling a beat. Musical entrainment occurs when internal oscillations synchronize to external, temporally regular or metric musical events. In music, interbeat intervals are usually isochronous or close to isochronous (see Honing, Bouwer, & Háden, 2014, for a review of musical beat perception and temporal regularity). Musical rhythms also conform to hierarchical metric structures.

Entrainment has implications beyond simply subserving the human ability to tap or dance along with a musical beat. Even nonhuman animals entrain to music (Patel, Iversen, Bregman, & Schulz, 2009; Schachner, Brady, Pepperberg, & Hauser, 2009). Entrainment capacities appear

early in human development and shape the relationship between infants and caregivers (Phillips-Silver & Keller, 2012). Entrainment drives the synchronization of musical performers (Zamm, Pfordresher, & Palmer, 2015), the relationship between communication, gesture, and music (Møller, Odell-Miller, & Wigram, 2002), and the development of individual identity within social music-making groups (Blacking, 1977). Entrainment also enables social bonding among groups, cultural rites of passage, and the weaving of people into communities (Lomax, 1982). Furthermore, entrainment can help patients with movement disorders (Thaut, McIntosh, & Hoemberg, 2015), modulate behavior in autistic patients (Orr, Myles, & Carlson, 1998), help autistic patients learn complex tasks (Møller, Odell-Miller, & Wigram, 2002), stabilize body rocking behavior in retarded adolescents (Rider & Eagle, 1986), help manage postoperative pain (Bradt, 2010), and help patients suffering from strokes and brain injuries (Thaut, Kenyon, Hurt, McIntosh, & Hoemberg, 2002).

Behavioral evidence suggests endogenous oscillations entrain to both isochronous sine tone sequences (Barnes & Jones, 2000; McAuley & Jones, 2003) and metric—although nonisochronous—sine tone sequences (Ellis & Jones, 2010; Large & Jones, 1999). Electroencephalographic evidence suggests that rhythmic musical stimuli entrain neural oscillations, and not merely superpositions of transient evoked responses (Henry, 2017). The importance of demonstrating entrainment effects in more ecologically valid, complex musical stimuli is becoming increasingly recognized. For instance, Tierney and Kraus (2013) demonstrated musical entrainment effects using pre-recorded popular music. These researchers emphasized that although simple sine tones afford tight experimental control, examining the temporal processing of complex musical stimuli is critical for understanding how people perceive music in everyday life.

Entrainment in speech and music. It is unknown whether the temporal aspects of music perception and speech perception are the same (Kotz & Schwartz, 2010), and the extent to which the rhythmic expectancies that drive musical rhythm perception are similar or identical to those that drive speech rhythm perception is a matter of continued debate (Beier & Ferreira, 2018). Furthermore, it is unknown whether rhythm tracking in music and speech rely on shared neural processes, a question to which an answer would likely help people with rhythmic and language impairments (Tierney & Kraus, 2013). It is possible that rhythm processing in speech and music share a biological basis (Peele & Davis, 2012). Recent reviews have linked music and language temporal processing by arguing they should recruit overlapping brain oscillations (Haegens & Golumbic, 2017), examined the likelihood that entrainment similarly drives the perception of non-isochronous speech rhythm and music rhythm (Beier & Ferreira, 2018), and made theoretical arguments in support of common music-language temporal processing mechanisms (Dilley, McAuley, & Dilley, 2012).

Shared resources. The shared resources view asserts that music and speech rhythm perception rely on shared neural and cognitive resources. It hypothesizes that music and speech rhythm perception share processing resources originating in the sensorimotor network that includes the premotor and supplementary motor areas, the cerebellum, and the basal ganglia (Kotz, Gunter, & Wonneberger, 2005; Kotz & Schmidt-Kassow, 2015). It also proposes that precise neural timing processes in the auditory system are dually responsible for auditory-motor entrainment and phonological skills (Tierney & Kraus, 2014). Evidence from an fMRI study suggests music and speech processing share anatomical resources but merely access them differently (Abrams, Bhatara, Ryali, Balaban, Levitin, & Menon, 2010).

Temporal processing abilities crucial to both speech and music perception are likely present at birth (Háden, Honing, Török, & Winkler, 2015). Temporal processing abilities in both music and speech in 9-month olds improved following 12-session musical interventions (Zhao & Kuhl 2016). Moreover, rhythm discrimination ability was positively related to sentence-in-noise—but not words-in-noise—perception (Slater & Kraus, 2016), and a 2-year longitudinal study showed musical training that included rhythm training improved speech-in-noise perception in low-income, bilingual elementary school students (Slater et al., 2015). Both musical rhythm and pitch processing skills predicted speech-pitch discrimination performance after controlling for relevant cognitive abilities (Morrill, McAuley, Dilley, & Hambrick, 2015). Neural oscillations entrain to periodicities in both music and speech, even when those periodicities are not truly isochronous (Hawkins, 2014). Moreover, people have been shown to synchronize their speech productions to both intelligible and unintelligible speech; non-linguistic factors such as pitch contour, amplitude envelope, and spectral stimulus properties were important for synchronization (Cummins, 2009).

A variety of studies have demonstrated beneficial transfer effects from musical rhythm training to language outcomes such as phonological, grammar, syntax, and literacy skills. For example, 5- to 8-year olds were more accurate judging whether aurally presented sentences were grammatically correct when the sentences followed regular compared to irregular 32 s percussion-timbre prime sequences (Chern, Tillmann, Vaughan, & Gordon, 2017). This grammaticality judgment facilitation effect was more robust in children with specific language impairments than in children with dyslexia (Przybylski et al., 2013). Nonetheless, children with dyslexia improved in phonological awareness skills after both a phoneme discrimination intervention that is commercially available and a newly developed intervention that highlights

the auditory rhythmic properties of speech and nonspeech sounds (Thomson, Leong, & Goswami, 2013). Performance on grammaticality sentence judgment tasks improved in 4- to 6-year olds following both musical and linguistic 20-day training programs (Janus, Lee, Moreno, & Bialystok, 2016). Syntax comprehension was better after (compared to before) eight morphosyntactic training exercise sessions when primed with both 30 s metric musical percussion-timbre sequences and 30 s environmental sound clips; grammaticality judgment accuracy and non-word repetition performance, however, was better only following the training sessions that were primed with music, and not environmental sounds (Bedoin, Besombes, Escande, Dumont, Lalitte, & Tillmann, 2017). Furthermore, musical rhythm discrimination ability accounted for over half of the variance in morpho-syntactic competence in typical language-developing children, even after controlling for prior musical activities, socioeconomic status, and non-verbal IQ (Gordon, Shivers, Wieland, Kotz, Yoder, & McAuley, 2014).

Studies using priming paradigms to investigate music-language cross-domain entrainment effects have been informative but limited in number and scope. One electroencephalographic priming study showed that isochronous sine-tone prime sequences decrease reaction times in phoneme-detection tasks (Cason & Schön, 2012). This study is limited in its ability to determine whether music and speech rhythm perception rely on the same processing resources, however, because it used pseudowords instead of real words, and sine tones instead of ecologically valid music. A similar study also demonstrated decreased phoneme-detection task reaction times to percussion-timbre rhythm primes (Cason, Astésano, & Schön, 2016), but this study examined the effects of rhythm independent of regularity and meter—each prime was a single rhythmic pattern, rather than a sequence of repeated patterns.

Independent resources. The independent resources view asserts that music and speech rhythm perception rely on independent neural, perceptual, and cognitive resources. One MEG study found that intelligible speech produced different patterns of correlated magnetic steady-state field activation across brain areas than unintelligible speech and music (Härle et al., 2004). Zoefel, Archer-Boyd, and Davis (2017) found that neural entrainment had a different influence on neural responses to intelligible speech stimuli than unintelligible speech. Moreover, systematically manipulating the phase relation between speech rhythm and neural oscillations using combined fMRI and TMS affected neural responses to intelligible—but not unintelligible—speech. Other results suggest that grouping linguistic and non-linguistic information relies on different neural mechanisms (Schmidt-Kassow, Rothermich, Schwartze, & Kotz, 2011).

Swaminathan and Schellenberg (2016, 2018) disagree with the interpretation of many correlational, and even some longitudinal, studies that music training causes improved nonmusical perceptual and cognitive abilities. These cognitive abilities include reading, vocabulary, phonological perception, linguistic stress processing, speech segmentation, speech intonation perception, second language learning, and verbal short-term, long-term, and working memory, in addition to more general abilities such as intelligence, IQ, and executive functions (see for review Schellenberg & Weiss, 2013). Schellenberg (2015) emphasizes the role of gene-environment interactions in associations between music training and nonmusical abilities. Demographic, personality, and cognitive variables are correlated with childhood musical training, and genetics plays a large role in music achievement and aptitude. Thus, it is possible that people with good nonmusical skills and personalities high in the openness Big Five personality trait seek out music training, instead of vice versa. Indeed, the correlation between

music training and nonverbal intelligence disappeared when music aptitude was controlled for, and the correlation between music aptitude and nonverbal intelligence stayed strong even when music training and socioeconomic status were controlled for (Swaminathan, Schellenberg, & Khalil, 2017). Moreover, the correlation between music training and reading ability (speed and comprehension) disappeared when controlling for general cognitive abilities (short-term memory, working memory, and nonverbal intelligence; Swaminathan, Schellenberg, & Venkatesan, 2018). One recent study even suggests that general cognitive ability and openness, in addition to music training, predict musical competence (the ability to perceive, discriminate, and remember beat or tone sequences; Swaminathan & Schellenberg, 2018).

Mixed results further call into question the premise that music, language, and general cognitive abilities share processing resources and benefit from transfer effects. Apart from the finding that musicians performed better than nonmusicians on tests of perceptual speed and flexibility of closure, musicians performed no differently than nonmusicians on tests of word fluency, verbal comprehension, memory, reasoning, number, and space (Helmbold, Rammsayer & Altenmüller, 2005). Moreover, despite the finding that musicians received higher verbal memory scores and nonmusicians, nonmusicians received higher intelligence scores than musicians (Brandler & Rammsayer, 2003). Finally, a 20-day children training program in either visual art or music produced no more improvement in phonological awareness skills in the music condition than in the visual art condition, despite the fact that auditory-visual skills (the ability to pair words with visual symbols) improved more in the music condition than in the visual art condition (Moreno, Friesen, & Bialystok, 2011).

The Present Research

Music and language share the fundamental attribute of unfolding dynamically in time. Despite the research on and implications of cross-domain music-language transfer effects, the extent of overlap between the neural and cognitive resources that underlie temporal information processing in music and speech remains unknown. Studies comparing the perception of music rhythm and speech rhythm have only recently begun to emerge (Kraus & Slater, 2015; Patel, 2008), and it is unclear whether shared neural and cognitive resources underlie music and speech entrainment (Kraus & Slater, 2015; Large & Jones, 1999; Tierney & Kraus, 2013, 2014). I aimed to provide evidence that helps elucidate the nature of this potential overlap. In the present research, I tested whether music-induced oscillations affect speech tempo discrimination performance, and whether speech-induced oscillations affect music tempo discrimination performance.

Seminal studies using behavioral priming paradigms to study entrainment (e.g., Barnes & Jones, 2000; Large & Jones, 1999; McAuley & Jones, 2003; McAuley & Kidd, 1998) have demonstrated that task performance is better when the ending time of the standard occurs in-phase with the prime periodicity than when the ending time of the standard occurs too early or too late relative to the prime periodicity. These results are illustrated as an inverted-U expectancy profile (see Figure 1). The accepted interpretation of the inverted-U expectancy profile is that the internal oscillations that become synchronized with the external prime sequence are the same oscillations involved in the perception of, and responsible for the performance benefits received by, subsequent test stimuli. By this reasoning, if the inverted-U expectancy profile emerges in the present research when the prime sequence is music and the standard-comparison pair is speech, then that would suggest the oscillations that became synchronized with the music were

the same oscillations involved in the perception of, and responsible for the performance benefits received by, the subsequent speech.

The shared-resources view predicts the inverted-U expectancy profile will emerge in both the within-domain conditions (music prime stimuli followed by music test stimuli; speech prime stimuli followed by speech test stimuli) and the between-domain conditions (music-speech; speech-music). The independent-resources view, on the other hand, predicts the inverted-U expectancy profile will emerge only in the within-domain conditions. If the mechanisms underlying music and speech temporal processing are independent, then 1) music- and speech-induced oscillations should affect music and speech perception, respectively, but 2) speech- and music-induced oscillations should not affect music and speech perception, respectively.

Method

Participants. Twenty-three undergraduate students with normal hearing enrolled in General Psychology at the University of Arkansas (10 females) ranging from years 18 to 20 ($M = 18.65$; $SD = 0.84$) received course credit for their participation. Sample size was determined by setting alpha to .05, power to .95, and effect size to .1 based on previous research (Barnes & Jones, 2000; Large & Jones, 1999; McAuley & Jones, 2003; GPower 3.1). None of the participants were music majors; seven reported having some level of musical training, but this consisted mostly of participation in school choir and band. The experiment was approved by the University of Arkansas Institutional Review Board, and all the participants gave informed consent before participating.

Design. The experiment had a fully crossed $3 \times 3 \times 2 \times 2$ mixed-factorial, fully within-subjects design: three prime tempi, three comparison tempi, two prime domains, and two test domains. The three prime tempi were either 20% slower than, faster than, or the same tempo as

the standard tempo. The three comparison tempi were either 12% slower than, faster than, or the same tempo as the standard tempo. The two prime domains were either music or speech. The two test domains were either music or speech.

Stimuli. Each trial consisted of a prime sequence, standard, and comparison (always in that order; see Figure 1). The prime sequence comprised seven repetitions of either a speech utterance or a musical phrase. The standard-comparison pair comprised either two identical speech utterances or two identical musical phrases. The onsets of both the standard and comparison always occurred in phase with the prime periodicity (the tempo manipulation affected only their offsets).

A silent interval separated the offset of the final (seventh) prime stimulus and onset of the standard. The duration of this silent interval was identical to the duration of one of the seven prime stimuli. A silent interval also separated the offset of the standard and the onset of the comparison. The duration of this silent interval varied as a function of the prime's tempo, such that the onset of the comparison was always in phase (on the beat) with the prime periodicity. The standard was always the same tempo. The standard tempo was 192 events (syllables or musical chords) per minute, or on average, one event every 313 ms. Specifically, the total duration of the standard was 1250 ms, the first three events (e.g., “pow-dered-with” for the speech utterance) were each 208 ms in duration, and the final event (e.g., “gold”) was 625 ms in duration.

For example, for a trial in which the duration of the prime utterance was 1000 ms, the silent interval separating the offset of the final prime stimulus and onset of the standard was 1000 ms, and the silent interval separating the offset of the standard and the onset of the comparison was 750 ms. For a trial in which the duration of the prime utterance was 1500 ms, the silent

interval separating the offset of the final prime stimulus and onset of the standard was 1500 ms, and the silent interval separating the offset of the standard and the onset of the comparison was 1250 ms.

Speech stimulus. The speech utterance was a man speaking the phrase “powdered with gold.” It was a modified version of a stimulus used in Tierney, Dick, Deutsch, and Sereno (2012). Tierney et al. collected a set of 24 utterances from audiobooks spoken by three males in equal proportions. Tierney et al. found behavioral and neurological evidence that the utterances are perceived as speech, not song, and that they do not transform into song following repetition. Moreover, I collected pilot data from eight participants who completed the present study, and the data showed that even when listening to “powdered with gold” in the current experiment, the utterance did not transform into song.

I made minor adjustments to the “powdered with gold” utterance used in Tierney et al. (2012) using Audacity (2.0.6). I created perfect interstress regularity, or speech stresses separated by equal durations, by increasing or decreasing the tempo of individual syllables. Furthermore, I acoustically manipulated the intervals between all (both stressed and unstressed) syllables to create purely metric utterances, ensuring that syllable onsets conformed to the 6:8 hierarchical metric structure. The duration of the standard was always 1250 ms.

For example, for the utterance “powdered with gold,” the duration between the onsets of the stressed syllables “pow-” and “gold” was identical to the duration between the onsets of the stressed syllables “gold” and the subsequent “pow-”. Furthermore, all four syllables (“pow-dered-with-gold”) conform to a 6:8 meter: The first three syllables “pow,” “dered,” and “with,” are equal in duration to three isochronous musical eighth notes. The fourth syllable “gold” is

equal in duration to a musical dotted quarter note; a dotted quarter note has the same duration as three consecutive eighth notes.

Music stimulus. I created a musical analog of the “powdered with gold” utterance by digitally manipulating the pitch contour (pattern of changes in pitch direction) of a segment of the first movement of Beethoven’s Fifth Symphony to match the pitch contour of the “powdered with gold” utterance (changes in mean pitch frequency between each syllable; determined using Praat 6.0.30). The pitch contour of both the musical analog and the speech utterance was: D#3 (155 Hz) – D#3 (155 Hz) – C#3 (139 Hz) – G#2 (104 Hz). The acoustic profile of the musical analog was different from that of the speech utterance only in timbre and pitch stability (the degree to which pitch frequency changed within each syllable or tone). The musical phrase and speech utterance had the same timing properties (overall duration, interonset interval duration, stress pattern, rhythm, and meter) and amplitude. Holding constant all acoustic variables except timbre and pitch stability ensured that potential cross-domain entrainment effects would not be attributable to extraneous acoustic features that covaried with the prime- and test-domain manipulations.

Procedure. Participants were tested individually. They sat in a 4’ x 4’ WhisperRoom MDL 4848E/ENV Sound Isolation Enclosure, faced a 22” Dell P2212H monitor, wore Sennheiser HD 600 headphones, and made responses using a computer keyboard and mouse. The auditory stimuli were presented binaurally at a comfortable listening level (approximately 65 dB). The experiment was presented on a Dell OptiPlex 7010 desktop computer running Windows 7. The experiment concluded with a brief demographic questionnaire and lasted about 1 hr.

Each experimental session comprised 12 blocks of nine trials (randomization occurred within and between blocks). All nine combinations of the standard-comparison tempo

manipulation occurred in each block. Over two counterbalanced experimental sessions on separate days, each participant heard all nine combinations of the standard-comparison tempo manipulations 12 times for each of the four prime-domain by test-domain combinations, totaling 432 trials.

In the first experimental session, participants heard either the speech prime-speech test condition or the speech prime-music test condition in randomized order. In the second experimental session, participants heard either the music prime-music test condition or the music prime-speech test condition in randomized order. We did this to avoid people perceiving the speech as song. If people had heard any of the music prime conditions before a speech prime condition, then it is possible that participants would have heard the speech utterances as more musical than they would have otherwise. This is because the musical prime condition might have induced a musical mode of listening. Moreover, if participants had heard any of the music prime conditions first, then any subsequent speech prime conditions would be affected by participants' memory of the music primes. Our pilot data together with Tierney's et al. (2012) categorization data show the speech stimuli in the present research were perceived as speech and not song.

Every trial in every condition involved the same tempo discrimination task: Participants chose whether the comparison was slower than, faster than, or the same tempo as the standard. Participants were required to respond within 2500 ms using the computer keyboard (slower = left arrow; same = down arrow; faster = right arrow). Trials in which participants did not respond within 2500 ms ended with the visual prompt, "Please respond quicker," and were discarded from data analysis (164 trials, 2% of the total 8468 trials, were discarded).

Before beginning, participants read on-screen instructions and studied a visual diagram of the stimuli and task. Participants then completed three practice trials that contained only the

standard-comparison pairs (no primes). Participants then completed an additional set of six practice trials that included the prime sequences. Participants were told to always ignore the prime sequences and focus only on the tempi of the standard and comparison.

Data Analysis

We ran a linear mixed model (Baayen, Davidson, & Bates, 2008) to account for random effects between participants. The random effect was Subject, and the within-subjects fixed effects were Prime Tempo (slower than, faster than, or the same tempo as the standard), Prime Domain (music or speech), and Test Domain (music or speech). Three participants did not respond within the time limit on more than 10% of their trials; their data were excluded from the analysis. The remaining participants exceeded the time limit on a total of 164 trials; these trials were excluded from analysis. The remaining data consisted of 8304 tempo judgments (either slower, faster, or the same as responses). We ran the model with restricted maximum likelihood, using the *lmer* function of the *lme4* package (Bates, Maechler, Bolker, & Walker, 2015) in R (R Core Team, 2015). We obtained regression weights using the *summary* function of the *lme4* package, *F* statistics and *p* values for omnibus effects (Satterthwaite approximation) using the *anova* function of the *car* and *lmerTest* packages (Fox & Weisberg, 2010), and normed means, standard deviations, and standard errors using the *summarySEwithin* function of the *Rmisc* package (Morey, 2008).

Results

Figure 3 shows the Prime Tempo by Prime Domain by Test Domain interaction, $F(2, 8273) = 5.66, p < .01$. This interaction qualified the Prime Domain by Test Domain interaction, $F(1, 8273) = 27.60, p < .0001$, the main effect of Test Domain, $F(1, 8286) = 10.51, p = .001$, and the main effect of Prime Tempo, $F(2, 8273) = 47.79, p < .0001$. The main effect of Prime Tempo

demonstrates an overall inverted U-shaped expectancy profile, a well-established measure of entrainment (McAuley & Jones, 2003). The inverted U-shaped expectancy profile suggests that the oscillations entrained by the prime stimuli affected the perception of the test stimuli—that shared resources underlie the perception of the prime and test. Therefore, the emergence of an inverted U-shaped expectancy profile in a between-domain condition would suggest the presence of shared resources for processing music and speech.

The Prime Tempo by Prime Domain by Test Domain interaction shows that the speech-speech inverted U-shaped expectancy profile was more robust than the speech-music inverted U-shaped expectancy profile, but that there was no significant difference between the music-music and music-speech inverted U-shaped expectancy profiles. Specifically, on the one hand, the reduction in accuracy between the same and slower prime-tempo conditions was significantly greater when both the prime and test were speech than when the prime was speech and the test was music, $\beta = 0.068$, $t(8273) = 2.86$, $SE = 0.024$, $p < .01$. Likewise, the reduction in accuracy between the same and *faster* prime-tempo conditions was significantly greater when both the prime and test were speech than when the prime was speech and the test was music, $\beta = 0.060$, $t(8273) = 2.52$, $SE = 0.024$, $p = .012$. On the other hand, the reduction in accuracy between the same and slower prime-tempo conditions was not significantly greater when both the prime and test were music than when the prime was music and the test was speech, $\beta = 0.024$, $t(8273) = 1.01$, $SE = 0.024$, $p = .312$. Similarly, the reduction in accuracy between the same and *faster* prime-tempo conditions was not significantly greater when both the prime and test were music than when the prime was music and the test was speech, $\beta = 0.043$, $t(8273) = 1.81$, $SE = 0.024$, $p = .070$.

Overall, the present results support the shared resources view more than the independent resources view. If completely independent resources were used to process speech and music rhythm, then there should have been no differences between any of the prime-tempo conditions in the between-domain conditions. This is not what we found. Furthermore, although the difference in accuracy between the same and slower prime-tempo conditions in the between-domain conditions was not significant, the difference in accuracy between the same and *faster* prime-tempo conditions in the between-domain conditions was significant. People were more accurate when the prime tempo was the same as the standard tempo than when the prime tempo was faster than the standard tempo both when the prime was speech and the test was music, $\beta = -0.053$, $t(8273) = -1.08$, $SE = 0.017$, $p < .01$, and when the prime was music and the test was speech, $\beta = -0.058$, $t(8273) = -3.53$, $SE = 0.017$, $p < .001$.

The difference between the inverted U-shaped expectancy profile in the music-prime–speech-test condition and the music-prime–music-test condition was not significant. This makes the evidence for the shared resources view stronger in the music-speech condition (where the resources that subserved the perception of speech were previously entrained by music) than in the speech-music condition. This is consistent with other research demonstrating that between-domain entrainment effects are more likely to emerge when people are primed with music than with speech (Dalla Bella, Palmer, & Jungers, 2003; Dickerson, 2012; Jungers, Hupp, & Dickerson, 2016; Simchy-Gross & Margulis, 2018). This might occur because music is inherently more temporally periodic and metric than speech. This explanation, however, cannot account for the present results because the music and speech in the present research were both equally periodic and metric. An alternative explanation is that music induces a musical mode of listening that engenders refined temporal perception. This explanation received support from

research that used the speech-to-song illusion to induce musical modes of listening and showed improved temporal regularity discrimination when people listened to speech utterances that sound more like song than speech (Graber, Simchy-Gross, & Margulis, 2017).

Discussion

Music and speech share many temporal features, but it is unclear the extent to which the resources that *process* music and speech are shared. Are the neural, perceptual, and attentional oscillations that synchronize with music stimuli the same oscillations that synchronize with metric speech stimuli? The present results suggest some, but not all, of the same oscillations synchronize with metric music and speech.

In the present research, people listened to metric prime stimuli and then compared the tempo of test stimuli. The prime stimuli were either music or speech, and the test stimuli were either music or speech. The prime stimuli were used to synchronize the listener's pulses of neural and attentional oscillations with the pulse of the prime meter. If music primes influence the temporal perception of subsequent metric speech stimuli (if an inverted U-shaped expectancy profile emerges), then that would suggest the oscillations that synchronize with music are the same oscillations that synchronize with metric speech. Likewise, if metric speech primes influence the temporal perception of subsequent music stimuli, then that would suggest the oscillations that synchronize with metric speech are the same oscillations that synchronize with music.

We found both that metric music primes affected the temporal perception of subsequent metric speech stimuli, and that metric speech primes affected the temporal perception of subsequent metric music stimuli. An inverted U-shaped expectancy profile emerged in both cases, providing evidence supporting the shared resources view. These between-domain

expectancy profiles, however, were less robust than the within-domain expectancy profiles. This suggests that some, but not all, of the resources that underlie speech and music perception are shared. More research is needed to determine the degree to which music and speech rhythm perception resources are shared.

Prime tempo asymmetry. One interesting aspect of the present results is the minor asymmetry represented in the inverted U-shaped expectancy profiles in the between-domain conditions (see Figure 3). In these conditions, the difference between the same and faster prime-tempo conditions was significant, but the difference between the same and slower prime-tempo conditions was not. People were significantly less accurate when the comparison was faster than the standard than when the comparison was the same tempo as the standard. In contrast, people were not significantly less accurate when the comparison was slower than the standard than when the comparison was the same tempo as standard. This result is interesting because both the slower and faster stimuli were different from the same stimulus by exactly 12%.

One reason the slower prime tempo conditions might have been easier than the faster prime tempo conditions is that changes in tempo perception are not linear. This explanation fits with research on time perception demonstrating parabolic relationships between actual and perceived duration. In the context of tempo, an example of this phenomenon is that the difference between the perception of four beats per second and two beats per second is not the same as the difference between the perception of eight beats per second and six beats per second. Further in line with this potential explanation for the asymmetry in the present findings is the fact that the faster prime tempo was 20% faster than the standard tempo, but the standard tempo was not 20% faster than the slower tempo—rather, the slower tempo was 20% slower than the standard tempo. These values are not equivalent. To demonstrate this, consider percentage

increases and decreases above and below the number 10. 50% less than 10 is 5. 50% more than 10 is 15. But 50% less than 15 is not 10—it is 7.5. The same principle might apply to tempo manipulations. If so, future research should calculate the slower tempo by answering the question, “What tempo is the prime 20% faster than?,” instead of “What tempo is 20% slower than the prime?”

Previous research using the present priming paradigm shows the degree to which the inverted U-shaped expectancy profile emerges depends on the degree to which the prime sequence disrupts the listener’s oscillator. Importantly, the expectancy profile found in previous research is continuous, not categorical. Previous studies using five, instead of three, prime tempo manipulations show (1) the expectancy profile is not always perfectly symmetrical and (2) differences in accuracy increase as the prime tempo moves away from the standard tempo (Barnes & Jones, 2000). Thus, in the present research, if we had made the slower prime tempo even slower, then it is likely the difference in accuracy between the slower and same prime tempo conditions would have been significant in both the within- and cross-domain conditions. Importantly, this would only make the slower trials more difficult if shared temporal processing resources subserve music and speech rhythm perception. If these resources were completely independent between music and speech, then this manipulation making the slower trials even slower should not affect task difficulty, and thus should not result in an inverted U-shaped expectancy profile.

Although the asymmetry in the present research was unanticipated, the asymmetry complements previous research. McAuley and Kidd (1988) presented people with two auditory tone sequences (a standard-comparison pair separated by a silent interval) and asked people to judge whether the comparison was slower, faster, or the same tempo as the standard. People were

more accurate (had lower tempo-discrimination thresholds) judging faster tempos in conditions where the comparison occurred too early, and people were more accurate judging slower tempos in conditions where the comparison occurred too late. These findings suggest not only that early onsets lead to perceptions of shorter durations between the event preceding an early onset and the early-onset event, itself, but also that early onsets lead to perceptions of shorter durations between the subsequent events in a sequence (the ones that generate the perception of tempo). The present research showed that people were more accurate (had more tempo-discrimination correct responses) judging standards that were faster than the preceding primes. The present findings thus suggest that it is easier to judge the difference in tempo between two ecologically valid multi-onset patterns when the tempo of the first pattern is faster the tempo of a preceding prime sequence than when the tempo of the first pattern is slower than the tempo of a preceding prime sequence.

Future research can account for the asymmetry in the present research by making the slower prime tempo greater than 12% slower than the standard. This would decrease the tempo of the prime (which entrains the listener's internal oscillator), and thus make the tempo difference between the listener's internal oscillator and the standard greater. The purpose of this adjustment would be to make the task more difficult and the difference between the slower and same prime-tempo conditions significant, assuming that shared resources underlie music and speech rhythm perception. Finding a significant difference between the slower and same prime-tempo conditions in the between-domain conditions would strengthen the evidence found in the present research and further suggest that shared resources are involved in the perception of rhythm in music and metric speech.

Other consideration. One aspect of the present research was the speech was purely regular and metric. Indeed, regularity of the prime in the present priming paradigm is required for the paradigm to be informative and useful. Nonetheless, the perfect regularity the speech primes limits the extent to which the current findings generalize to conversational dialogue. Natural spoken dialogue, although rhythmic, is rarely purely metric. Thus, the present findings generalize more to poetic speech and chanting than to everyday spoken conversations.

One limitation of the present research is the use of only one utterance and one analogous musical phrase. In order to maximize that degree that the auditory stimuli in the present research generated robust entrainment, we used a single spoken utterance and its musical analog. The use of a single utterance conforming to a 6:8 meter, rather than multiple different utterances conforming to different meters, ensured that the oscillations of the listeners maintained a steady and robust pattern throughout the experiment. Future research can use multiple different utterances and musical analogues to increase the generalizability of the results. If the same pattern of results is found, it not only would replicate the present findings, but it also would suggest that shared resources underlie the perception of music and speech rhythm even when the strength of oscillatory patterns are less robust.

Implications

Ecological validity. One aspect of the present research that is significant to the literature on timing and entrainment is the finding that ecologically realistic music and speech stimuli (although metric) produced the inverted U-shaped expectancy profile found when using only sine tone sequences. Finding this profile using dynamic, ecologically valid stimuli is meaningful because the current time perception literature lacks an investigation of psychological time using complex, temporal, and ecologically valid stimuli (Matthews & Meck, 2014). This shortcoming

in the literature weakens the extent to which researchers can generalize many findings to more common situations in daily life, making decades of research less practical than it ought to be. The present study is the first study to the authors knowledge that used complex auditory patterns instead of simple sine tones to demonstrate the classic inverted U-shaped expectancy profile. This is significant because until now, researchers have not tested whether the entrainment effects produced by sine tone sequences generalize to more complex and ecologically valid stimuli. The present research suggests they do.

Future research is needed to establish the reliability and robustness of this generalization. The generalization from simple sine tones to complex music and speech patterns provides stronger evidence of entrainment than sine tone sequences alone. This is because sine tone prime sequences provide a clearer periodicity for oscillations to entrain to than music and speech patterns that have more continuous temporal profiles, such as the ones in the present research. The music and speech stimuli in the present research induced entrainment via sound amplitude modulations during continuous sound, in contrast to discrete sine tones separated by silences. Thus, the fact that the present findings showed inverted U-shaped expectancy profiles lends evidence to support the view that entrainment is pervasive and robust in everyday life.

Interval timing and rhythmic timing. Interval timing refers to the generation of explicit duration judgments, or consciously estimating, for example, the overall duration of a musical song or spoken conversation. Rhythmic timing, on the other hand, refers to the synchronization of internal neural, perceptual, and attentional oscillations with external rhythmic events, or implicitly processing the beat of your favorite song or recurrent inter-syllabic stress of your conversational partner's speech.

The present research demonstrates the utility of mechanisms underlying rhythmic timing, in contrast to interval timing. McAuley and Jones (2003) provided evidence supporting the view that the functional mechanisms underpinning time perception of temporal patterns are oscillatory and entrainable (e.g., dynamic attending theory; Large & Jones 1999) rather than discrete and switch-based (e.g., scalar expectancy theory; Gibbon, Church, & Meck, 1984). The present research adds additional evidence supporting this view by demonstrating inverted U-shaped expectancy profiles with stimuli comprising complex, multi-event patterns. These findings serve as evidence to discount explanations proposed by interval timing theories, such as the multiple look hypothesis (Drake & Botte, 1993).

The multiple look hypothesis posits that averaging in memory the duration between each event in a prime sequence makes the judgment of the duration between events in the test stimuli more accurate. If we had used isochronous sine tone sequences to entrain listeners' oscillations, then the multiple look hypothesis would have been able to account for the present findings. In that case, the multiple look hypothesis would have argued that the inverted U-shaped expectancy profiles emerged because listeners recorded a running average in reference memory of the duration between each successive tone in the isochronous prime sequence, and then used this average duration as a reference when comparing the duration between two test intervals. Thus, accuracy would be highest when the standard interval equaled that average. This explanation, however, cannot explain the present findings. This is because the auditory stimuli used in the present research were multi-event patterns that were not isochronous (they were nonisochronous and metric), and because listeners judged the *tempo* of more than two events ("powdered with gold" comprises four events, or onsets) in the test stimuli rather than the duration between only two events.

Music training and language processing. The present research exists within a body of work that seeks to understand the relationship between musical training and language perception. This work suggests that overlapping neural resources subserve the perception of music and speech pitch and timbre, in addition to rhythm. Proposals have been developed to explain why musical training might improve the neural processing of speech (Besson, Chobert, & Marie, 2011; Patel, 2011, 2014; Strait & Kraus, 2011) to account for growing correlational, experimental, and longitudinal evidence. Evidence suggests that musicians have more refined auditory brainstem speech encoding than nonmusicians (Kraus & Chandrasekaran, 2010), that the strength of this pattern increases as the amount of musical training increases (Musacchia, Sams, Skoe, & Kraus, 2007), that auditory brainstem responses improve in quality even after relatively short, eight-week training sessions (Song, Skoe, Wong, & Kraus, 2008), that musical training provided across longitude studies leads to improved auditory cortical function and structure (Hyde et al., 2009), and that this improved auditory brainstem speech encoding correlates with real-world speech perception abilities (Bidelman, Gandour, & Krishnan, 2011). These real-world speech perception abilities include consonant perception (Elmer, Hänggi, Meyer, & Jäncke, 2013), syllable perception and phonological processing (Chobert, François, Velay, & Besson, 2012), prosody perception (Parbery-Clark, Skoe, & Kraus, 2009; Thompson, Schellenberg, & Husain, 2004), and hearing in noise and reading ability (Banai et al., 2009).

The present findings provide evidence for the first component of the OPERA hypothesis (Patel, 2011, 2014)—that premise music and speech rhythm perception rely on overlapping neural networks. Moreover, the OPERA hypothesis would predict that listening to and performing the types of musical phrases used as primes in the present research repeatedly relatively long periods of time with focused attention and emotional engagement (as often occurs

with formal musical training) would result in between-domain, experience dependent auditory brainstem plasticity, or improvements in auditory brainstem speech encoding, speech perception, and reading ability.

One interesting aspect of the design of the present research that has been largely untested is the role of attention in the present priming paradigm. In the present research, participants were instructed to ignore the prime sequence (following all previous research using this priming paradigm). An interesting avenue for future research is the investigation of how changing these instructions to *encourage* attention to the prime sequence in between-domain conditions might influence the results. Two assertions of the opera hypothesis—attention and memory—predict that the inverted U-shaped expectancy profile would be strengthened, and that the difference between the slower-faster and the same prime tempo conditions would become greater. The opera hypothesis asserts that music-language transfer affects strengthen as the attention to which people pay to music while listening to and performing music increases, and as the emotional engagement with which people experience during music listening and performing increases. Therefore, if instructing people to attend to the prime sequence in the present priming paradigm increases attention to and emotion experienced during the music prime, then the music-language transfer effects (the inverted U-shaped expectancy profile) should become more robust. This finding would both support the opera hypothesis and provide additional evidence suggesting that overlapping neural networks are involved in the perception of rhythm and metric music and speech.

Conclusion

We used novel, ecologically realistic stimuli in an established priming paradigm to examine whether shared or independent resources subserve the perception of rhythm in metric

music and speech. Music-induced oscillations influenced the temporal perception of subsequent speech stimuli, and vice versa, suggesting at least some of the resources subserving music and speech rhythm perception are shared. One promising direction for future research would be studying overlapping neural networks between music and speech tempo processing with the present priming paradigm using neurological measures, such as EEG and MEG. Collecting neural, in addition to behavioral, measures is critical for establishing between-domain auditory brainstem plasticity. More broadly, to conclusively establish between-domain neural plasticity from music training with purely instrumental music to language processing, each individual future study must include behavioral, neural, and longitudinal data before and after music training.

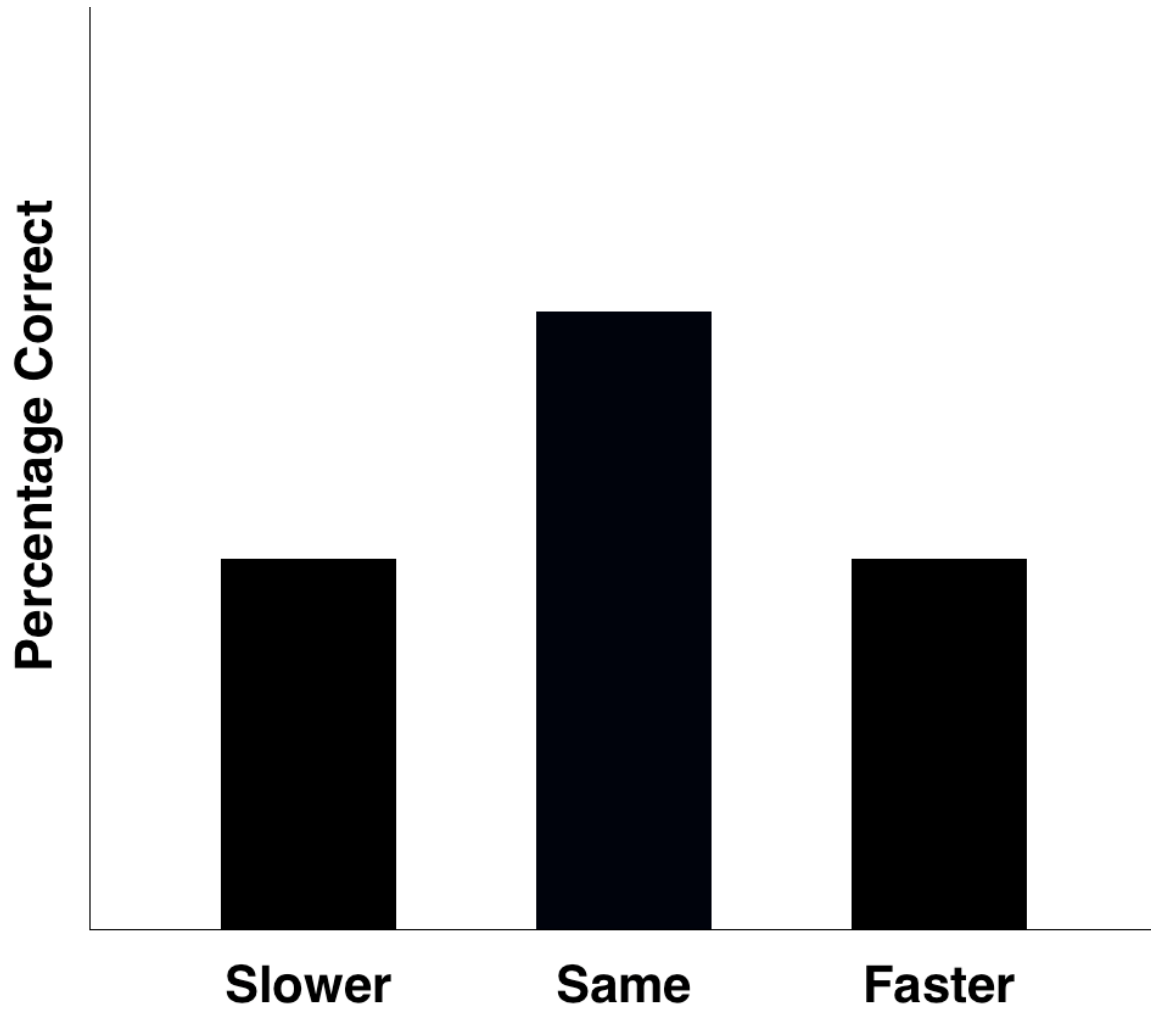
Figures

Figure 1. Sample inverted U-shaped expectancy profile.

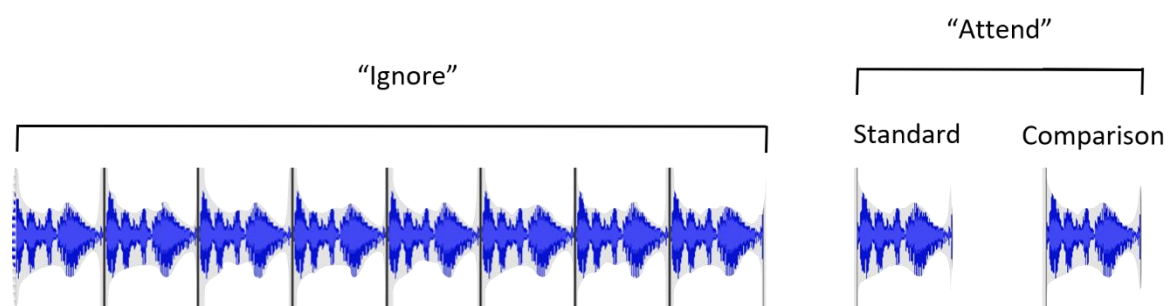


Figure 2. Prime sequence followed by standard and comparison pair, all of the same tempo.

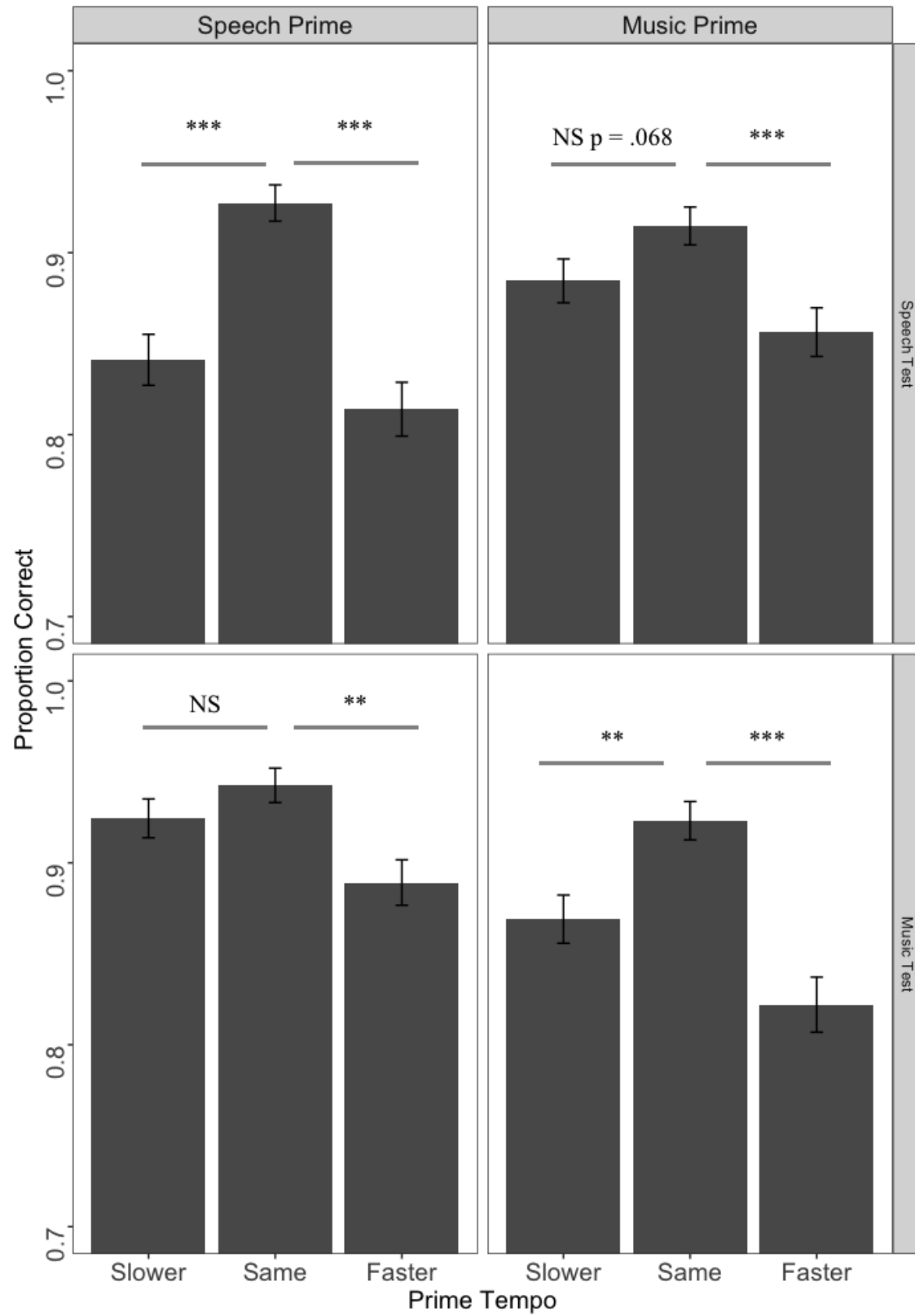


Figure 3. Proportion Correct as a function of Prime Tempo, Prime Domain, and Test Domain.

References

- Abrams, D. A., Bhatara, A., Ryali, S., Balaban, E., Levitin, D. J., & Menon, V. (2010). Decoding temporal structure in music and speech relies on shared brain resources but elicits different fine-scale spatial patterns. *Cerebral Cortex*, *21*, 1507-1518.
- Ammirante, P., Patel, A. D., & Russo, F. A. (2016). Synchronizing to auditory and tactile metronomes: a test of the auditory-motor enhancement hypothesis. *Psychonomic Bulletin & Review*, *23*, 1882-1890.
- Banai, K., Hornickel, J., Skoe, E., Nicol, T., Zecker, S., & Kraus, N. (2009). Reading and subcortical auditory function. *Cerebral Cortex*, *19*, 2699-2707.
- Barnes, R., & Jones, M. R. (2000). Expectancy, attention, and time. *Cognitive Psychology*, *41*, 254-311.
- Bedoin, N., Besombes, A. M., Escande, E., Dumont, A., Lalitte, P., & Tillmann, B. (2017). Boosting syntax training with temporally regular musical primes in children with cochlear implants. *Annals of Physical and Rehabilitation Medicine*. doi:10.1016/j.rehab.2017.03.004.
- Beier, E. J., & Ferreira, F. (2018). The temporal prediction of stress in speech and its relation to musical beat perception. *Frontiers in Psychology*, *9*, 431. doi:10.3389/fpsyg.2018.00431
- Besson, M., Chobert, J., & Marie, C. (2011). Transfer of training between music and speech: Common processing, attention, and memory. *Frontiers in Psychology*, *2*, 94. <http://dx.doi.org/10.3389/fpsyg.2011.00094>.
- Bhide, A., Power, A., & Goswami, U. (2013). A rhythmic musical intervention for poor readers: A comparison of efficacy with a letter-based intervention. *Mind, Brain, and Education*, *7*, 113-123.
- Bidelman, G. M., Gandour, J. T., & Krishnan, A. (2011). Cross-domain effects of music and language experience on the representation of pitch in the human auditory brainstem. *Journal of Cognitive Neuroscience*, *23*, 425-434.
- Blacking, J. (1977). *The anthropology of the body*. London: Academic Press.
- Boersma, Paul & Weenink, David (2018). Praat: doing phonetics by computer [Computer Program]. Version 6.0.42, retrieved from <http://www.praat.org/>
- Bradt, J. (2010). The effects of music entrainment on postoperative pain perception in pediatric patients. *Music and Medicine*, *2*, 150-157.
- Brandler, S., & Rammsayer, T. H. (2003). Differences in mental abilities between musicians and non-musicians. *Psychology of Music*, *31*, 123-138.

- Buzsáki, G., & Draguhn, A. (2004). Neuronal oscillations in cortical networks. *Science*, 304, 1926-1929.
- Chobert, J., François, C., Velay, J. L., & Besson, M. (2012). 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, 956-967.
- Corriveau, K. H., & Goswami, U. (2009). Rhythmic motor entrainment in children with speech and language impairments: tapping to the beat. *Cortex*, 45, 119-130.
- Cummins, F. (2009). Rhythm as entrainment: The case of synchronous speech. *Journal of Phonetics*, 37, 16-28.
- Drake, C., & Botte, M. C. (1993). Tempo sensitivity in auditory sequences: Evidence for a multiple-look model. *Attention, Perception, & Psychophysics*, 54, 277-286.
- Elmer, S., Hänggi, J., Meyer, M., & Jäncke, L. (2013). Increased cortical surface area of the left planum temporale in musicians facilitates the categorization of phonetic and temporal speech sounds. *Cortex*, 49, 2812-2821.
- Gibbon, J., Church, R. M., & Meck, W. H. (1984). Scalar timing in memory. *Annals of the New York Academy of Sciences*, 423, 52-77.
- Giles, H., Coupland, N. & Coupland, J. (1992). Accommodation theory: Communication, context and consequences. In Giles, H., Coupland, J., & Coupland, N. (Eds.), *Contexts of accommodation: Developments in applied sociolinguistics* (pp. 1–68). New York: Cambridge University Press.
- Gordon R. L., Fehd H. M., McCandliss B. D. (2015). Does music training enhance literacy skills? a meta-analysis. *Frontiers in psychology*, 6.
- Gordon, R. L., Shivers, C. M., Wieland, E. A., Kotz, S. A., Yoder, P. J., & Devin McAuley, J. (2015). Musical rhythm discrimination explains individual differences in grammar skills in children. *Developmental Science*, 18, 635-644.
- Goswami, U. (2011). A temporal sampling framework for developmental dyslexia. *Trends in Cognitive Sciences*, 15, 3-10.
- Goswami, U. (2012). Language, music, and children's brains: A rhythmic timing perspective on language and music as cognitive systems. In Rebuschat, P., Rohrmeier, M., Hawkins, J. A., & Cross, I. (Eds.), *Language and music as cognitive systems* (pp. 292-301). Oxford, England: Oxford University Press.

- Graber, E., & Fujioka, T. (2017). *Actively anticipating upcoming tempo changes modulates induced neural beta power*. Paper presented at the 1st Annual Conference of the Timing Research Forum, Strasbourg, France.
- Háden, G. P., Honing, H., Török, M., & Winkler, I. (2015). Detecting the temporal structure of sound sequences in newborn infants. *International Journal of Psychophysiology*, 96, 23-28.
- Härle, M., Rockstroh, B. S., Keil, A., Wienbruch, C., & Elbert, T. R. (2004). Mapping the brain's orchestration during speech comprehension: task-specific facilitation of regional synchrony in neural networks. *BMC Neuroscience*, 5, 40.
- Haegens, S., & Golumbic, E. Z. (2017). Rhythmic facilitation of sensory processing: a critical review. *Neuroscience & Biobehavioral Reviews*, 86, 150-165.
- Hawkins, S. (2014). Situational influences on rhythmicity in speech, music, and their interaction. *Philosophical Transactions of the Royal Society B*, 369, 20130398.
- Helmbold, N., Rammsayer, T., & Altenmüller, E. (2005). Differences in primary mental abilities between musicians and nonmusicians. *Journal of Individual Differences*, 26, 74–85.
- Henry, M. J. (2017). *Separating stimulus-driven and entrained neural responses using musical rhythms*. Paper presented at the 1st Annual Conference of the Timing Research Forum, Strasbourg, France.
- Henry, M. J., Herrmann, B., & Obleser, J. (2016). Neural microstates govern perception of auditory input without rhythmic structure. *Journal of Neuroscience*, 36, 860-871.
- Herrmann, B., Henry, M. J., Haegens, S., & Obleser, J. (2016). Temporal expectations and neural amplitude fluctuations in auditory cortex interactively influence perception. *Neuroimage*, 124, 487-497.
- Honing, H., Bouwer, F. L., & Háden, G. P. (2014). Perceiving temporal regularity in music: The role of auditory event-related potentials (ERPs) in probing beat perception. In Merchant, H., & de Lafuente, V. (Eds.), *Neurobiology of interval timing* (pp. 305-323). New York, NY: Springer.
- Hutchins, E. (1995). *Cognition in the wild*. Cambridge, MA: MIT Press.
- 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, 3019-3025.
- Jarvis, B. G. (2014). DirectRT (Version 2014.1.127). New York, NY: Empirisoft Corporation.

- Jungers, M. K., & Hupp, J. M. (2009). Speech priming: Evidence for rate persistence in unscripted speech. *Language and Cognitive Processes*, 24, 611-624.
- Jungers, M. K., Hupp, J. M., & Dickerson, S. D. (2016). Language priming by music and speech. *Music Perception*, 34, 33-39.
- Kraus, N., & Chandrasekaran, B. (2010). Music training for the development of auditory skills. *Nature Reviews Neuroscience*, 11, 599-605.
- Kraus, N., & Slater, J. (2015). Music and language: relations and disconnections. In Celesia, G.G., Hickok, G. (Eds.), *The human auditory system: Fundamental organization and clinical disorders. Handbook of clinical neurology* (pp. 207–222). Amsterdam: Elsevier. doi:10.1016/b978-0-444-62630-1.00012-3
- Large, E. W., & Jones, M. R. (1999). The dynamics of attending: How people track time-varying events. *Psychological Review*, 106, 119.
- Lawrance, E. L., Harper, N. S., Cooke, J. E., & Schnupp, J. W. (2014). Temporal predictability enhances auditory detection. *The Journal of the Acoustical Society of America*, 135, EL357-EL363.
- Lomax, A. (1982). The cross-cultural variation of rhythmic style. In Davis, M. (Ed.), *Interaction rhythms: periodicity in communicative behavior* (pp. 149-174). New York: Human Sciences Press.
- Manson, J. H., Bryant, G. A., Gervais, M. M., & Kline, M. A. (2013). Convergence of speech rate in conversation predicts cooperation. *Evolution and Human Behavior*, 34, 419-426.
- Matthews, W. J., & Meck, W. H. (2014). Time perception: The bad news and the good. *Wiley Interdisciplinary Reviews: Cognitive Science*, 5, 429-446.
- McAuley, J. D., & Jones, M. R. (2003). Modeling effects of rhythmic context on perceived duration: A comparison of interval and entrainment approaches to short-interval timing. *Journal of Experimental Psychology: Human Perception and Performance*, 29, 1102.
- McAuley, J. D., & Kidd, G. R. (1998). Effect of deviations from temporal expectations on tempo discrimination of isochronous tone sequences. *Journal of Experimental Psychology: Human Perception and Performance*, 24, 1786.
- Møller, A. S., Odell-Miller, H., & Wigram, T. (2002). Indications in music therapy: Evidence from assessment that can identify the expectations of music therapy as a treatment for autistic spectrum disorder (ASD); meeting the challenge of evidence based practice. *British Journal of Music Therapy*, 16, 11-28.
- Moreno, S., Friesen, D., & Bialystok, E. (2011). Effect of music training on promoting preliterate skills: Preliminary causal evidence. *Music Perception*, 29, 165-172.

- Morrill, T. H., McAuley, J. D., Dilley, L. C., & Hambrick, D. Z. (2015). Individual differences in the perception of melodic contours and pitch-accent timing in speech: Support for domain-generalty of pitch processing. *Journal of Experimental Psychology: General*, 144, 730.
- 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*, 104, 15894-15898.
- Nozaradan, S., Peretz, I., Missal, M., & Mouraux, A. (2011). Tagging the neuronal entrainment to beat and meter. *Journal of Neuroscience*, 31, 10234-10240.
- Oever, T. (2017). *Temporal expectations influence entrainment presence and strength*. Paper presented at the 1st Annual Conference of the Timing Research Forum, Strasbourg, France.
- Orr, T. J., Myles, B. S., & Carlson, J. K. (1998). The impact of rhythmic entrainment on a person with autism. *Focus on Autism and Other Developmental Disabilities*, 13, 163-166.
- Parbery-Clark, A., Skoe, E., & Kraus, N. (2009). Musical experience limits the degradative effects of background noise on the neural processing of sound. *Journal of Neuroscience*, 29, 14100-14107.
- Patel, A. D. (2003). Language, music, syntax and the brain. *Nature Neuroscience*, 6, 674-681.
- Patel, A. D. (2008). *Music, language, and the brain*. New York: Oxford University Press.
- Patel, A. D. (2011). Why would musical training benefit the neural encoding of speech? The OPERA hypothesis. *Frontiers in Psychology*, 2.
- Patel, A. D. (2012). Language, music, and the brain: a resource-sharing framework. *Language and Music as Cognitive Systems*, 204-223.
- Patel, A. D. (2014). Can nonlinguistic musical training change the way the brain processes speech? The expanded OPERA hypothesis. *Hearing Research*, 308, 98-108.
- Patel, A. D., Iversen, J. R., Bregman, M. R., & Schulz, I. (2009). Studying synchronization to a musical beat in nonhuman animals. *Annals of the New York Academy of Sciences*, 1169, 459-469.
- Peretz, I., & Coltheart, M. (2003). Modularity of music processing. *Nature Neuroscience*, 6, 688-691.
- Phillips-Silver, J., & Keller, P. (2012). Searching for roots of entrainment and joint action in early musical interactions. *Frontiers in Human Neuroscience*, 6, 26.

- Przybylski, L., Bedoin, N., Krifi-Papoz, S., Herbillon, V., Roch, D., Léculier, L., ... & Tillmann, B. (2013). Rhythmic auditory stimulation influences syntactic processing in children with developmental language disorders. *Neuropsychology*, 27, 121.
- Rider, M. S., & Eagle Jr, C. T. (1986). Rhythmic entrainment as a mechanism for learning in music therapy. In Evans, J.R., & Clynes, M. (Eds.), *Rhythm in psychological, linguistic and musical processes* (pp. 225-249). Springfield, IL: C.C. Thomas.
- Rike, L. (2015). *The cross-domain priming of language and motor rate* (Doctoral dissertation). Retrieved from The Ohio State University.
- Schachner, A., Brady, T. F., Pepperberg, I. M., & Hauser, M. D. (2009). Spontaneous motor entrainment to music in multiple vocal mimicking species. *Current Biology*, 19, 831-836.
- Schellenberg, E. G. (2015). Music training and speech perception: a gene–environment interaction. *Annals of the New York Academy of Sciences*, 1337, 170-177.
- Schellenberg, E. G., & Weiss, M. W. (2013). Music and cognitive abilities. In Deutsch, D. (Ed.), *The Psychology of Music* (pp. 499–550). Amsterdam: Elsevier.
- Schmidt-Kassow, M., Rothermich, K., Schwartz, M., & Kotz, S. A. (2011). Did you get the beat? Late proficient French-German learners extract strong–weak patterns in tonal but not in linguistic sequences. *Neuroimage*, 54, 568-576.
- Semin, G. R., & Cacioppo, J. T. (2008). Grounding social cognition: Synchronization, coordination, and co-regulation. In Semin, G. R. & Smith, E. R. (Eds.), *Embodied grounding: Social, cognitive, affective, and neuroscientific approaches* (pp. 119–147). New York: Cambridge University Press.
- Slater, J., & Kraus, N. (2016). The role of rhythm in perceiving speech in noise: A comparison of percussionists, vocalists and non-musicians. *Cognitive Processing*, 17, 79-87.
- 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.
- Slevc, L. R., & Okada, B. M. (2015). Processing structure in language and music: a case for shared reliance on cognitive control. *Psychonomic Bulletin & Review*, 22, 637-652.
- Slevc, L.R., Rosenberg, J.C., & Patel, A.D. (2009). Making psycholinguistics musical: Self-paced reading time evidence for shared processing of linguistic and musical syntax. *Psychonomic Bulletin & Review*, 16, 374-381. doi:10.3758/16.2.374
- Song, J. H., Skoe, E., Wong, P. C., & Kraus, N. (2008). Plasticity in the adult human auditory brainstem following short-term linguistic training. *Journal of Cognitive Neuroscience*, 20, 1892-1902.

- Strait, D., & Kraus, N. (2011). Playing music for a smarter ear: Cognitive, perceptual and neurobiological evidence. *Music Perception*, 29, 133-146.
- 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, 191-201.
- Swaminathan, S., & Schellenberg, E. G. (2016). Music training. In Strobach, T. & Karbach, J. (Eds.), *Cognitive training* (pp. 137–144). Basel, Switzerland: Springer International Publishing. http://dx.doi.org/10.1007/978-3-319-42662-4_13
- Swaminathan, S., & Schellenberg, E. G. (2017). Musical competence and phoneme perception in a foreign language. *Psychonomic Bulletin & Review*, 1-6.
- Swaminathan, S., & Schellenberg, E.G. (2018). Music training and cognitive abilities: Associations, causes, and consequences. In Thaut, M. H. & Hodges, D. A. (Eds.), *The Oxford handbook of music and neuroscience*. New York: Oxford University Press. doi: 10.1093/oxfordhb/9780198804123.013.26
- Swaminathan, S., & Schellenberg, E.G. (2018). Musical competence is predicted by music training, cognitive abilities, and personality. *Scientific Reports*, 8:9223. doi: 10.1038/s41598-018-27571-2
- Swaminathan, S., Schellenberg, E.G., & Khalil, S. (2017). Revisiting the association between music lessons and intelligence: Training effects or music aptitude? *Intelligence*, 62, 119-124. doi: 10.1016/j.intell.2017.03.005
- Swaminathan, S., Schellenberg, E.G., & Venkatesan, K. (2018). Explaining the association between music training and reading in adults. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 44, 992-999. doi: 10.1037/xlm0000493
- Thaut, M. H., Kenyon, G. P., Hurt, C. P., McIntosh, G. C., & Hoemberg, V. (2002). Kinematic optimization of spatiotemporal patterns in paretic arm training with stroke patients. *Neuropsychologia*, 40, 1073-1081.
- Thaut, M. H., McIntosh, G. C., & Hoemberg, V. (2015). Neurobiological foundations of neurologic music therapy: Rhythmic entrainment and the motor system. *Frontiers in Psychology*, 5, 1185.
- Thompson, W.F., Schellenberg, E.G., & Husain, G. (2004). Decoding speech prosody: Do music lessons help? *Emotion*, 4, 46-64.
- Thomson, J. M., Leong, V., & Goswami, U. (2013). Auditory processing interventions and developmental dyslexia: A comparison of phonemic and rhythmic approaches. *Reading and Writing*, 26, 139–161.

- Tierney, A., Dick, F., Deutsch, D., & Sereno, M. (2012). Speech versus song: multiple pitch-sensitive areas revealed by a naturally occurring musical illusion. *Cerebral Cortex*, 23, 249-254.
- Tierney, A., & Kraus, N. (2013). Neural responses to sounds presented on and off the beat of ecologically valid music. *Frontiers in Systems Neuroscience*, 7.
- Tierney, A., & Kraus, N. (2014). Auditory-motor entrainment and phonological skills: precise auditory timing hypothesis (PATH). *Frontiers in Human Neuroscience*, 8.
- Trehub, S. E., & Hannon, E. E. (2006). Infant music perception: Domain-general or domain-specific mechanisms? *Cognition*, 100, 73-99.
- Zamm, A., Pfordresher, P. Q., & Palmer, C. (2015). Temporal coordination in joint music performance: effects of endogenous rhythms and auditory feedback. *Experimental Brain Research*, 233, 607-615.
- Zoefel, B., Archer-Boyd, A., & Davis, M.H. (2017). *Phase entrainment of neural oscillations is causally relevant for neural responses to intelligible speech*. Paper presented at the 1st Annual Conference of the Timing Research Forum, Strasbourg, France.

Appendix



Office of Research Compliance
Institutional Review Board

September 21, 2017

MEMORANDUM

TO: Rhimmon Simchy-Gross
Elizabeth Hellmuth Margulis

FROM: Ro Windwalker
IRB Coordinator

RE: EXEMPT PROJECT CONTINUATION

IRB Protocol #: 16-10-142

Protocol Title: *Music Language Persistence*

Review Type: ☒ EXEMPT

New Approval Date: 09/21/2017

Your request to extend the referenced protocol has been approved by the IRB. We will no longer be requiring continuing reviews for exempt protocols.

If you wish to make any modifications in the approved protocol that may affect the level of risk to your participants, you must seek approval *prior to* implementing those changes. All modifications should be requested in writing (email is acceptable) and must provide sufficient detail to assess the impact of the change.

If you have questions or need any assistance from the IRB, please contact me at 109 MLKG Building, 5-2208, or irb@uark.edu.