The Effects of Music on Auditory-Motor Integration for Speech:

A Behavioral Priming and Interference Study

by

Arianna LaCroix

A Thesis Presented in Partial Fulfillment
of the Requirements for the Degree
Master of Science

Approved March 2015 by the
Graduate Supervisory Committee:

Corianne Rogalsky, Chair
Shelley Gray
Julie Liss

ARIZONA STATE UNIVERSITY

May 2015

ABSTRACT

Language and music are fundamentally entwined within human culture. The two domains share similar properties including rhythm, acoustic complexity, and hierarchical structure. Although language and music have commonalities, abilities in these two domains have been found to dissociate after brain damage, leaving unanswered questions about their interconnectedness, including can one domain support the other when damage occurs? Evidence supporting this question exists for speech production. Musical pitch and rhythm are employed in Melodic Intonation Therapy to improve expressive language recovery, but little is known about the effects of music on the recovery of speech perception and receptive language. This research is one of the first to address the effects of music on speech perception. Two groups of participants, an older adult group (*n=24; M = 71.63 yrs*) and a younger adult group (*n=50; M = 21.88 yrs*) took part in the study. A native female speaker of Standard American English created four different types of stimuli including pseudoword sentences of normal speech, simultaneous music-speech, rhythmic speech, and music-primed speech. The stimuli were presented binaurally and participants were instructed to repeat what they heard following a 15 second time delay. Results were analyzed using standard parametric techniques. It was found that musical priming of speech, but not simultaneous synchronized music and speech, facilitated speech perception in both the younger adult and older adult groups. This effect may be driven by rhythmic information. The younger adults outperformed the older adults in all conditions. The speech perception task relied heavily on working memory, and there is a known working memory decline associated with aging. Thus, participants completed a working memory task to be used as a covariate in analyses of differences across stimulus types and age groups. Working memory ability was found to correlate with speech perception performance, but that the age-related performance differences are still significant once working memory differences are taken into account. These results provide new avenues for facilitating speech perception in stroke patients and sheds light upon the underlying mechanisms of Melodic Intonation Therapy for speech production.

ACKNOWLEDGEMENTS

I would like to express the deepest appreciation to my committee chair, Dr. Corianne Rogalsky, who offered continual encouragement, enthusiasm, and insight. Without her guidance, persistent help, and support, this thesis would not have been possible.

I would also like to thank Nicole Blumenstein for her help with creating and recoding stimuli. I am also grateful to Nicole for her assistance with the transcription and scoring of participant responses, as well as helping with countless other tasks. I cannot thank her enough for her frequent aid.

Lastly, I would like to acknowledge my grandparents, Alan and Diane Gotlieb, for introducing me to the “older adult group” and coordinating the data collection with each of these participants. Without them, the “older adult group” would not have existed.

This research was supported by the GRAMMY Foundation Scientific Research Grant (Principle Investigator: Corianne Rogalsky, PhD) and by Arizona State University.

TABLE OF CONTENTS

Page

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

Introduction

It has long been established that music and language have a number of commonalities. Language and music are both rule-governed systems that allow for an infinite number of messages to be transmitted based on a finite set of symbols (Trehub, 2003; Besson & Schön, 2001). These two domains also share properties of rhythm, acoustics, and hierarchical structures (Hausen, Torppa, Salmela, Vainio, & Sarkamo, 2013; Hannon, 2009; Patel & Daniele, 2003; Patel, Peretz, Tramo, & Labreque, 1998) and there is evidence of shared neural processes underlying music and language processing within neural computations (Patel, 2013; Zatorre & Baum, 2012; Jiang et al., 2012; Fedorenko, Behr, & Kanwisher, 2011; Fedorenko, Patel, Casasanto, Winawer, & Gibson, 2009; Loui, Alsop, & Schlaug 2009; Steinbeis & Koelsch, 2008; Mandell, Schulze, & Schlaug, 2007; Koelsch et al., 2004). Previous research suggests that speech and music processing are homologous, meaning the two systems have similar structures and anatomical positioning indicating a common evolutionary pathway. Brown (2001) argues that music and speech are homologous systems, utilizing both associated and dissociated processes that contain shared, parallel, and distinct features. Brown (2001) describes shared features as using the same neural computations, parallel features as using analogous neural computations, and distinct features as using separate neural computations in processing language and music. This view that speech and music are homologous systems provides a foundation for understanding the associations and dissociations of speech and music, and will help researchers to better understand how music may aid speech perception.

The association between language and music is more clearly understood in terms of expressive language. Clinical research has primarily focused on music's ability to aid expressive language recovery, but little research exists that explores the effects of music on speech perception and receptive language recovery. To date, the current literature on speech perception and music focuses on the neural correlates associated with speech, language, and music perception, as well as the effects of musical training programs on speech perception abilities. The present study differs as it solely looks at whether music has an effect on speech perception in the absence of a musical training program. Specifically this research aims to provide insight into the transfer of music onto speech perception in the non-musician brain. Thus, the purpose of the present study is to explore if music can aid speech perception in demanding situations and if so, what aspects of music are most beneficial in aiding speech perception. This thesis will also provide background information on the historical significance of language and music, the effects of music on expressive and receptive language, how musical training may affect the way the brain uses musical resources to interpret and express language, and the relationship between working memory, speech, and music.

Review of the Literature

Evolution and Development of Language and Music

Language and music are fundamental human characteristics that have an expansive history of societal interconnectedness. Both systems are utilized to convey human thoughts and emotions. These thoughts and emotions can be translated into meaningful messages that are communicated between individuals and convey warnings and group support (McMullen & Saffran, 2004; Trehub, 2003). These commonalities between the two domains may be exhibited in the evolutionary timeline with music possibly playing a role in the evolution of language (Masataka, 2007; Wallin, Merker, & Brown, 2000) although there is some disagreement (Berwick, Friederici, Chomsky, & Bolhuis, 2013). The link between musical vocalizations and speech vocalizations can be seen in the use of motherese or infant directed speech (IDS). A look at the use of IDS provides some insights into the evolution of language from music. IDS is a singsong manner that parents use to communicate with their infants. IDS is characterized by high pitch vocalizations that have an exaggerated rhythm and pitch contour. These singsong vocalizations that parents use when communicating with their infants are an intermediary step in the child's acquisition of language (Falk, 2004). Additionally, IDS also facilitates bonding between the infant and parent, and this facilitation of bonding in conjunction with promoting language acquisition is likely to be favored by natural selection as both are important traits that augment the infant's chances of survival (Falk, 2004).

Additional answers to how language evolved from music may lay in the relationship between tonal and atonal languages. Tonal languages use different pitches to distinguish between different phonemes and convey complex speech messages (Koelsch & Siebel, 2005). Tonal languages such as Mandarin use pitch to distinguish between individual words while atonal languages such as English use changes in pitch at the sentence level to convey an overall semantic meaning (i.e. a rising intonation signals asking a question) (Dediu & Ladd, 2007). Dediu and Ladd (2007) conducted research investigating whether there is a link between the emergence of atonal languages and gene mutations. The research concluded that a mutation of the ASPM gene might account for the split between atonal and tonal languages; people with the gene mutation are more likely to speak an atonal language while those without the mutation speak a tonal language (Dediu & Ladd, 2007). This research suggests that atonal languages may have evolved from tonal languages, which have more musicality incorporated within them, however more research needs to be conducted to substantiate this preliminary evidence.

There is also evidence that music played a role in communication, cooperation, and group cohesion in early societies, and that it continues to be a component of these contexts in modern day society (Koelsch & Siebel, 2005). Drums were one of the earliest musical instruments used to communicate over long distances. Talking drums are commonly used in Central and West Africa as a means of communication between villages. The Ewe tribe of West Africa is one such group that uses drums as a means of communication between villages (Brandt, Gebrian, & Slevc, 2012). The drummers use a combination of low and high tones, and pauses, to send messages to other villages and these messages may carry warnings of attack, invitations to ceremonies, or simply messages asking a friend to meet somewhere (Carrington, 1971). Tone is the key to communicating through drum language and this tonal counterpart of drum language is mimicked in the tonal nature of the majority of oral African languages (Carrington, 1971). Moving away from Africa, it is possible to see the continued interconnectedness of music and language in other cultures. On the Canary Islands people who speak the Silbo Gomero language utilize different whistles as their means of communication

(Brandt et al., 2012.) This informal observation of modern day cultures reveals the continued interconnectedness of music and language in the present day.

It is generally accepted that music and language have been linked throughout history, so recent research has focused on understanding the inner workings of both domains across the lifespan. Research on infants suggests that an infant's first use of language is heavily based on prosodic features. Infants begin to comprehend word boundaries and meanings through the prosodic aspects of language including melody and rhythm (Koelsch & Siebel, 2005). This research suggests that there is an innate aspect to musical perception, and research with pre-linguistic infants has corroborated the above findings demonstrating that infants have the same musical perception abilities as adults who have had continuous informal exposure to music for their entire lives, implying an inherent component of musical perception (Trehub, 2003).

Music is also hypothesized to play an important role in a child's development of emotions, social skills, and cognitive skills (Trehub, 2003). Research on neonates indicates that IDS and instrumental music engage overlapping neural networks (Kotilahti et al., 2010). Links have also been established between musical deficits, language impairments, and learning disabilities (Brandt et al., 2012). These correlations put forward the idea that the musical aspects of language and the overlapping neural computations for speech and music in neonates may be the foundation for the later development of the syntactic and semantic aspects of language (Kotilahti et al., 2010; Koelsch & Siebel, 2005). This research implies that speech and music are highly associated at birth; however, it is possible that as the child develops they dissociate. Case studies of adults who have amusia[1] without aphasia indicate some extent of dissociation between the domains of language and music. Peretz, Belleville, and Fontaine (1997) describe one such case in which the patient has intact oral language but is severely impaired in the musical domain, including her ability to process and express music. Case studies of individuals with amusia without aphasia suggest a dissociation between music and language in the brain (Piccirilli, Sciarma, & Luzzi, 2000; Dalla Bella, & Peretz, 1999; Griffiths et al., 1997; Peretz, Belleville, & Fontaine, 1997; Peretz et al., 1994; Peretz & Kolinsky, 1993; Peretz, 1990), but these cases are rare and thus may not reflect how the computations of music and language are related in most individuals.

Neurobiology of Language and Music

Evidence from hierarchical structure. While music and language clearly share acoustic and computational properties, little is understood about how they are related within the structure of the brain. Research exists that supports both the association and dissociation of the neural correlates of music and language. The neuroimaging data is mixed and points to both an association and dissociation of speech and music processing while the neuropsychology data points solely to a dissociation of the two (Patel, 2013; Zatorre & Baum, 2012; Patel, 2005; Patel, 2003; Patel, et al., 1998). Neuroimaging results from work by Rogalsky, Rong, Saberi, and Hickok (2011) specify dissociation between the neural computations engaged during the perception of language and music. They highlight the superior temporal lobe being substantially activated by both linguistic

---

[1] Amusia is a disorder resulting in impairments in pitch processing, music memory and recognition. Amusia can be either acquired, which results from a brain lesion, or congenital, which is present at birth with unknown etiology (Pearce, 2004).

and melodic stimuli, however, the specific areas of activation within the superior temporal lobe are different for each type of stimulus with piano melodies activating the dorsomedial temporal lobe extending into the insula and sentences more ventrolateral regions of the temporal lobe (Rogalsky et al., 2011). Similar results were reported by Abrams et al. (2011). When comparing speech and melodic stimuli to temporally scrambled versions (i.e. such that the speech and melodies had their temporal structure disrupted in an identical way), Abrams et al. (2011) revealed that speech and melodies activated similar neural computations, however, the manner in which these neural computations were activated and used differed, which corroborates Rogalsky et al.'s (2011) findings. Using the temporally scrambled music and speech, Abrams et al. (2011) established a dissociation between the brain's response to the temporal information of speech and music even though the two modes were transformed (i.e. scrambled) in an identical manner. The fact that temporal information is processed in a domain-specific way suggests that the dissociation of music and language in terms of processing structure and syntax may be related to differences between temporal processing for the two domains

Fedorenko et al. (2011) also report that language networks are distinct from musical processing networks. Sentences were found to activate a large left hemisphere fronto-temporo-parietal network, as seen in numerous other studies of language. Fedorenko et al. (2011) then examined how these language networks responded to music. Melodic stimuli were found to activate similar regions but to a lesser degree, however, the activation did not survive correction for multiple comparisons (Fedorenko et al., 2011). A caveat to interpreting Fedorenko et al.'s (2011) results lies in the presentation method of the stimuli; the sentences were visually presented while the music stimuli were presented auditorally. The difference of input modality may be confounding as reading is known to activate overlapping but distinct networks compared to speech perception (Price, 2012). However, their results still do support dissociation between the processing of language and music.

Evidence from pitch. Several pieces of evidence from studies of pitch processing suggest substantial overlap of the neural computations needed to support language and music processing (Patel, 2013; Zatorre & Baum, 2012; Jiang et al., 2012; Fedorenko, Patel, Casasanto, Winawer, & Gibson, 2009; Loui, Alsop, & Schlaug 2009; Steinbeis & Koelsch, 2008; Mandell, Schulze, & Schlaug, 2007; Koelsch et al., 2004). Evidence from congenital amusia research highlights the similarities in the neural computations utilized in processing music and language. Neuroimaging data indicates that individuals with amusia have bilateral structural abnormalities in the temporal and frontal lobes, including in regions long implicated in the acoustic processing of speech (Mandell, Schulze, & Schlaug, 2007). Loui, Alsop, and Schlaug (2009) also identified a reduced level of connectivity in the arcuate fasciculus, which connects regions known to support sensory-motor integration for speech and phonological processing (Hickok & Poeppel, 2007).

There is also behavioral evidence that individuals with congenital amusia have deficits in contour pitch processing for speech as well as deficits in fine-grained pitch processing, which is necessary for music and suggests some overlap of the neural computations for processing pitch (Patel, 2013). For example, individuals with congenital amusia have reduced capacity to distinguish between questions and statements when they differ solely by falling or rising intonation (Patel, 2013). Contour pitch processing is responsible for the interpretation of linguistic intonation, as it is highly sensitive to the broad-spectrum (sentence level) pitch shifts characteristic of language (Patel, 2013). These deficits in contour pitch processing are in addition to deficits in fine-grained pitch processing. Fine-grained pitch processing is primarily used to process music, as it is more precise in processing the rapid pitch changes characteristic of music (Patel, 2013). Research with individuals with congenital amusia demonstrate additional deficits in fine-grained pitch processing as evidenced by their inability to identify a dissonant note in a given melody (Peretz & Hyde, 2003).

Evidence from electroencephalography (EEG) studies also supports this notion that amusics have structural abnormalities and demonstrate abnormal brain activity during speech comprehension tasks. Jiang et al. (2012) used EEG to highlight the inability of amusics to correctly classify prosody as appropriate or inappropriate while listening to speech. Using the P600 and N100 event-related potentials (ERPs), Jiang et al. (2012) highlighted the relatively equal responses of the P600 and N100 while amusics processed both appropriate and inappropriate intonation, and this pattern differed considerably from the processing responses identified in healthy controls. The healthy controls had a larger P600 and smaller N100 response when processing inappropriate intonation and a smaller P600 and larger N100 response when processing appropriate intonation (Jiang et al., 2012). The abnormal ERPs and structural anomalies of known language regions suggest that amusics have pitch-processing deficits in both domains indicating an overlap of the neural computations required for processing pitch changes in both linguistic and melodic stimuli (Patel, 2013).

Some research with healthy controls supports the aforementioned results. Control subjects were found to process melodic stimuli within the superior temporal plane, including areas anterior and posterior to Heschl's gyrus, as well as regions of the middle and superior temporal gyri (Fedorenko, McDermott, & Kanwisher, 2012). These regions are consistent with classical areas of activation for auditory language (Rogalsky et al., 2011; Fedorenko et al., 2011; Price, 2010; Fedorenko et al., 2009) implying an association of the neural computations of language and music.

Evidence from rhythm. Rhythm is another commonality between music and language that has been investigated using neuroanatomical techniques. Rhythm consists of the grouping and segmenting of words or notes to create meaningful phrases. Phrase boundaries are marked by rate slowing in both music and language and also by pre-boundary lengthening of phonemes in language (Patel et al., 1998). Patel and Daniele (2003) investigated the link between the rhythm of music and the rhythm of language. The researchers compared the rhythm of classical music and language in English and in French and found that the composition's melodic rhythm resembled the linguistic rhythm of the composer's native language (Patel & Daniele, 2003) suggesting a link between the processing and expression of musical and linguistic rhythm. Listeners are also able to perceive the parallels of rhythm in speech and music across various cultures. Hannon (2009) found that both musician and non-musician listeners are able to identify which language an instrumental composition belongs to based on the melody's rhythm indicating that there is an association in the processing of linguistic and musical rhythm. In a review of the literature, Slevc (2012) highlights that speech rhythm and melodic rhythm are interconnected and that both utilize bilateral temporal processing. While temporal lobe structures in general play a role in both the processing of speech and music, there is ample evidence that these two types of stimuli utilize overlapping but distinct neural computations (Rogalsky et al., 2011; Fedorenko et al., 2011; Abrams et al., 2011). The distinct, yet close proximity of anatomical resources within the temporal lobe for linguistic and melodic rhythm processes may account for conflicting findings regarding shared versus distinct resources for speech and music.

Neurobiology: Theoretical implications. The seemingly conflicting neuroimaging, behavioral, and patient evidence across structures used for pitch and rhythm supports Patel's (2005) hypothesis that music and language dissociate in terms of long-term storage but share the same neural computational resources. This hypothesis indicates that damage to the long-term storage area of music or language, within the temporal regions of the brain, would result in deficits specific to that domain, while damage to the frontal lobe where the activation pathways (that stimuli follow once sorted by the auditory cortex) are located would result in deficits to both domains (Patel, 2013). In support of this hypothesis, Patel (2005) found that some aphasics who have damage to the frontal lobe have syntactic processing deficits in music and language, where as those with intact frontal regions and damage to the temporal lobe solely experience deficits in the affected domain, suggesting possible shared processing pathways within the left frontal lobe. It is likely that the possible overlap of language and melodic neural computations is due to common task demands, such as working memory, or similar input and output measures, with auditory stimuli as input and vocalizations as output (Zatorre & Baum, 2012). The contradiction in findings between lesion studies and neuroimaging studies also may be explained by a resource sharing framework (Patel, 2012) or the domains having homologous anatomy with shared, parallel, and distinct neural computations (Brown, 2001). Regardless of the computations driving a possible overlap, similarities between the language and music processing domains indicates that the recruitment of neural computational resources of musical processing may aid speech perception in patients when classical language networks are damaged.

Implications of Working Memory for Language and Music

It is well documented that language and working memory are closely related (Makuuchi, Bahlmann, Anwander, & Friederici, 2009; Archibald & Gathercole, 2006; Baddeley, 2003; Nation, Adams, Bowyer-Crane, & Snowling, 1999). Working memory is the ability to hold and manipulate information in immediate memory (Baddeley, 2003). This is important as working memory allows for reanalysis and comprehension of degraded or structurally complex speech (Rogalsky & Hickok, 2011; Rogalsky, Matchin, & Hickok, 2008; Wingfield & Grossman, 2006; Just, Carpenter, Keller, Eddy, & Thulborn, 1996; King & Just, 1991). Working memory is also critical in language development. The act of learning a new word is a basic working memory task that requires one to hear the word and hold it in working memory while extracting its meaning from the surrounding context (Ellis & Sinclair, 1996). As we age, our working memory capacity declines (Gazzaley, Cooney, Rissman, & D'Esposito, 2005). Raz, Rodrigue, Head, Kennedy, and Acker (1997) demonstrated that reduced relative gray matter brain volume correlates strongly with age in the prefrontal cortex, which, in addition to inferior parietal cortex, is the neural substrate supporting working memory. Age-related differences also include focal activation patterns in young adults as compared to more extensive patterns of activation in older adults during a variety of language and memory tasks (Wingfield & Grossman, 2006). The diffuse activation pattern in older adults has been suggested to reflect the need for additional cognitive effort and attention resources (Wingfield & Grossman, 2006). This is interesting because although working memory functioning declines, language comprehension remains stable (Wingfield & Grossman, 2006). The stability of language comprehension may suggest that other neural computations are being recruited as working memory capacity decreases. Exactly what these other neural computations are and how they contribute to the stability of language comprehension is not well understood.

While verbal working memory has been studied extensively, the idea of tonal working memory is less understood. Furthermore, the relationship between verbal and tonal working memory is understood even less. Research exists supporting both a convergence and divergence of verbal and tonal working memory networks (Burunat, Alluri, Toiviainen, Numminen, & Brattico, 2014; Semal, Demany, Ueda, & Halle, 1996; Salame & Baddeley, 1989). It has been documented that tonal working memory has a right hemisphere dominant lateralization while verbal working memory has a left hemispheric dominant activation pattern (Burunat et al., 2014). Salame and Baddeley (1989) discovered that vocal music interfered more with a verbal working memory task than instrumental music, which indicates a separate activation loop for both verbal and tonal working memory. Conversely, Semal et al. (1996) identified that interference stimuli, tones or phonemes, that were similar in pitch to the target stimuli had greater interference effects than tones or phonemes that were more dissimilar. This research indicates that pitch similarity is more important than modality in working memory tasks and suggests one working memory loop for both tonal and verbal working memory.

The neuroimaging literature indicates a convergence of the verbal and tonal working memory loops. Research in non-musicians identifies similar neural computations being involved in both verbal and tonal working memory (Schulze & Koelsch, 2012; Schulze, Zysset, Mueller, Friederici, & Koelsch, 2011; Koelsch, Schulze, Sammler, Fritz, Müller, & Gruber, 2009; Hickok, Buchsbaum, Humphries, & Muftuler, 2003). Both verbal and tone computations recruit a large fronto-parietal circuit including pre-motor cortex, anterior insula, supramarginal gyrus, intraparietal sulcus, planum temporale, inferior frontal gyrus, pre-supplementary motor area, and the cerebellum (Schulze & Koelsch, 2012). Additionally, verbal working memory tasks activated additional areas including the mid-dorsolateral prefrontal cortex, left inferior frontal sulcus, right ventral pre-motor, left intraparietal sulcus, right inferior parietal lobe, and the right cerebellum (Schulze et al., 2011). In overlapping regions for the verbal and tonal working memory tasks, the activations are stronger for verbal working memory, suggesting that tonal working memory recruits resources that are tuned to speech (Koelsch et al., 2009). This finding is corroborated with behavioral studies, which reveal higher performance on verbal tasks compared to tonal tasks in non-musicians (Schulze et al., 2011; Koelsch et al., 2009).

Musicians showed a more extensive pattern of activation than non-musicians for both verbal and tonal working memory (Schulze & Koelsch, 2012; Schulze et al., 2011; Koelsch et al., 2009). In musicians, activation of Broca's area, left pre-motor cortex, left insular cortex, pre-supplementary motor area, cingulate gyrus, and left inferior parietal lobe were all activated more strongly in tonal working memory tasks than verbal working memory tasks (Schulze et al., 2011). This is the opposite of what was found in non-musicians who showed greater activation of these areas during verbal working memory and suggests an overlap of the two systems. Additionally, musicians showed patterns of activation exclusive to verbal and tonal working memory. The specific activations included the right insular cortex for verbal working memory and basal ganglia, and left cerebellum for tonal working memory (Schulze et al., 2011). These activation patterns provide evidence for homologous systems of working memory, meaning verbal and tonal working memory share similar resources as well as have distinct resources. This understanding of the association and dissociation of tonal and verbal working memory is crucial to understanding whether music has an effect on speech perception.

Expressive Language and Music

The relationship between expressive language and music has been the focus of numerous research studies assessing expressive language recovery in patients with aphasia, particularly regarding the effects musical pitch and rhythm may have on expressive language recovery (Zumbansen, Peretz, Hébert, 2014; Lim et al., 2013;

Conklyn, Novak, Boissy, Bethoux, & Chemali, 2012; Stahl, Kotz, Henseler, Turner, & Geyer, 2011; Norton, Zipse, Marchina, & Schlaug, 2009; Schlaug, Marchina, & Norton, 2009; Schlaug, Marchina, & Norton, 2008; Belin et al., 1996). The interactions of language, rhythm and pitch, discussed above, have led to interventions that use music to recruit residual neural computations in both the dominant and non-dominant hemispheres (Conklyn et al., 2012; Schlaug et al., 2008). Imaging studies conducted on lesion patients reveal two main pathways of language recovery post stroke. These two pathways of language rehabilitation are the result of traditional aphasia therapy focusing on auditory comprehension and verbal expression (Chapey, 2010). These rehabilitation pathways include increased activation of the perilesional cortex in the left hemisphere and activation of the homologous language network of the right hemisphere (or non-dominant hemisphere) (Schlaug et al., 2008). Evidence from fMRI studies of aphasics with large left hemisphere lesions demonstrates substantial activation of language homologues in the right hemisphere after individuals participated in Melodic Intonation Therapy (MIT) (Conklyn et al., 2012). This increased activation of the right language homologues corresponded to improved expressive language skills (Conklyn et al., 2012). This research indicates that the brain's natural response to language recovery is to recruit new neural computations to facilitate the recovery of expressive language in individuals with large left hemisphere lesions post stroke.

The dorsal-ventral model of speech processing (Hickok & Poeppel 2000) provides a framework for understanding language reorganization and aphasia treatments. Briefly, the dorsal-ventral model includes two pathways, (1) a left-lateralized dorsal stream, which is utilized for speech production, repetition, and some phonological encoding, and (2) a bilateral ventral stream, which supports speech comprehension (Hickok & Poeppel, 2000). While both streams are activated in response to auditory stimulation, the dominant pathway is dependent upon the nature of the task. The ventral stream facilitates the processing of semantic information (Hickok & Poeppel, 2000) while the dorsal stream includes resources in the inferior parietal and frontal lobes to access sub-lexical phonological representations of speech (Hickok & Poeppel, 2000). Repetition solely activates the dorsal stream, as individuals do not need to comprehend the linguistic stimulus in order to repeat it (Hickok & Poeppel, 2007). During language recovery, the recruitment of right-hemisphere homologues of the typically left-lateralized dorsal stream suggests that networks outside of typical language areas may be able to compute sensory-motor interactions for speech.

Melodic Intonation Therapy (MIT) is a clinical intervention that may capitalize on this possible dorsal stream functional reorganization, and more generally the brain’s natural response to use undamaged homologous structures to attempt to compensate for expressive language deficits (Lim et al., 2013; Conklyn et al, 2012). MIT claims to engage the use of intact non-dominant hemisphere functions, including singing, to facilitate the recovery of expressive language in non-fluent aphasics, however research on its efficacy is still inconclusive to date (Conklyn et al., 2012; Schlaug et al., 2008). The treatment method involves the melodic intoning and rhythmic tapping of the left hand to produce words/phrases of increasing length (Norton et al., 2009). The overall nature of the task, repetition of the target phrase/sentence (Norton et al., 2009), relies on the dorsal pathway previously discussed (Hickok & Poeppel, 2007).

There are three levels of MIT: elementary, intermediate, and advanced. The specific procedure used by therapists varies from session to session depending on individual patient needs; there are no strict procedures for the intervention (Norton, et al., 2009) but the following steps are generally incorporated across interventionists. There are five steps in the elementary level procedure. In stage one, the therapist begins by showing the patient a visual cue of the phrases while humming the melody of the phrase one time at a rate of one syllable per second. Then the therapist intones the phrase two times while tapping the patient's left hand one time per syllable. Traditionally, MIT involves singing and humming two pitches with the higher pitch being assigned to the accented syllable and the lower pitch to the unaccented syllable, but current practices show more variability in the number of pitches utilized (Norton et al., 2009). Stage two of the elementary level consists of the therapist and patient intoning the target phrases together while the therapist taps the patient's left hand one time per syllable. In phase three, the therapist and patient start out intoning together and then halfway through the therapist fades their vocal support until the patient is singing independently with the sole accompaniment of the left hand-tapping. In stage four, the therapist intones and taps the target phrase while the patient listens, then the patient immediately taps and intones the same phrase. Once the patient has successfully completed stage four, the therapist sings a question and the patient responds by intoning the target phrase (Norton et al., 2009). MIT can either be concluded at the elementary level, or the patient can progress on to the intermediate level.

There are four stages in the intermediate level procedure and each stage is similar to the elementary level procedure except for the addition of the delayed repetition step and deletion of the humming step. Humming is deleted in step one, so step one of the intermediate level is identical to step one in the elementary level minus the humming component. The delayed repetition step takes the place of the immediate repetition step in the elementary procedure. In the delayed repetition phase the therapist intones and taps the target phrase while the patient listens and then after a six second delay the patient intones the phrase only using left hand tapping as a support (Norton et al., 2009). Successful completion of the intermediate level can either conclude MIT or the patient can move on to the advanced level procedure.

There are five steps in the advanced level MIT procedure. Step one is the delayed repetition stage and follows the same procedures as outlined in the intermediate protocol. Step two begins with the therapist introducing Sprechgesang to the patient. Sprechgesang is not intoning, but is instead speech production with an exaggerated emphasis on rhythm and syllable stress (Norton et al., 2009). In step two, the therapist presents the target phrase two times in Sprechgesang while tapping patient's left hand. In step three, the therapist and patient begin to speak the phrase in Sprechgesang but the therapist fades their support until the patient independently completes the phrase with only the support of left hand tapping. In stage four, the therapist presents the target phrase with normal intonation and no hand tapping, and after a six second delay the patient repeats the phrase with normal speech prosody. Step five begins six seconds after successful completion of step four at which time the therapist asks a question using normal prosody to elicit the target phrase. The patient responds to the question by saying the target phrase with normal prosody and no other assistance (i.e. no hand tapping) (Norton et al., 2009). The goal of stage five is to transition the patient to the rhythm and stresses of natural speech.

While the neurobiology of MIT is not well researched, clinical evidence promotes the efficacy of MIT for the recovery of expressive language in non-fluent aphasia (Lim et al., 2013; Conklyn et al., 2012; Stahl et al., 2011; Norton et al., 2009; Schlaug et al., 2008). It is thought that the musical intonation and rhythmic tapping of the left hand are the components of music that stimulate the functions of the right hemisphere (Lim et al., 2013). Schlaug et al. (2008) proposes that melodic intoning is responsible for the initial engagement of the right temporal lobe during MIT and that tapping the left hand acts as a priming mechanism for the right sensorimotor network. Interestingly, speech-language pathologists and music therapists identify singing as the critical factor in MIT, but little research exists to support this claim (Schlaug et al., 2008). However, research by Stahl et al. (2011) highlights rhythm as the crucial factor increasing expressive language production. Rhythm is an underlying component of music and may explain why therapists claim intoning to be the mitigating factor of MIT as well as provide support for Schlaug et al.'s (2008) proposal that intoning and left hand tapping, which both employ aspects of rhythm, are the significant components of MIT.

As previously stated, the neurobiology underlying MIT is not well-researched and the research that does exists is contradictory and inconclusive in regards to how intact neural computations are being accessed to promote the recovery of expressive language. Some evidence exists that suggests that MIT increases activity in the left hemispheres of

Broca's aphasics who participate in therapy while other evidence points to the recruitment of the right hemisphere in the processing of linguistic stimuli (Schlaug et al., 2008). This recruitment of right hemisphere function may indicate that music is able to facilitate activation of dorsal-ventral pathways in the right hemisphere and subsequently aid speech perception (Hickok & Poeppel, 2007). In reality, a combination of the two previous proposals is probable. It is likely that the right hemisphere structures recruited to aid in language processing are at increased levels of activation as compared to healthy controls if there is widespread left hemisphere damage, or in the acute stage of smaller lesions in left hemisphere (Conklyn et al., 2012).

Case studies provide some insight into the reorganization of language structures as a consequence of specific interventions following stroke. Schlaug et al. (2008) conducted a study that compared the effects of MIT on language recovery versus the effects of a control intervention, non-intoned speech therapy. The participants included two patients with non-fluent aphasia as a result of a left hemisphere stroke. The patients were chosen because they had lesions of approximately the same size and location. Differences in activation patterns between pre and post intervention neuroimaging indicated that the patient who received MIT therapy showed increased activation of the right hemisphere, including pre-motor frontal, inferior frontal, and temporal regions, while the patient who received non-intoned speech therapy revealed changes in activation primarily in the left hemisphere, namely the pre-central, post central, and superior temporal gyri (Schlaug et al., 2008).

Two other prominent studies exist that show similar results in terms of activation of the right hemisphere. Belin et al. (1996) used positron emission tomography (PET) scans to investigate changes in the brain during MIT therapy. Their results revealed increased activation of the right hemisphere and of Broca's area when patients underwent MIT therapy compared to patients whose therapy involved solely speech training (Belin et al., 1996). Schlaug et al.'s (2009) research uncovered similar findings using diffusion tensor imaging. Their research highlighted increased white matter density in the arcuate fasciculus in the right hemisphere when MIT was utilized (Schlaug et al., 2009). While Belin et al.'s (1996) and Schlaug et al.'s (2009) imaging results implicate right hemisphere involvement for increasing expressive language output, no studies to date have investigated whether music, and rhythm in particular, may facilitate receptive language recovery.

Links between music and speech production have also been investigated in other populations including the effects of singing and music on second language acquisition. Recent research conducted by Ludke, Ferreira, and Overy (2014) investigated whether singing could facilitate the acquisition of a foreign language. This study is the first experimental research of its kind to establish a link between singing and foreign language acquisition. Participants were assigned to three conditions: singing, speaking with normal prosody, and rhythmic speaking. The participants listened to novel phrases in an unfamiliar language, Hungarian, for a 15 minute training period after which they were tested using a variety of Hungarian language tests. Participants assigned to the singing condition demonstrated increased performance on the Hungarian language tests compared to participants in the normal speaking and rhythmic speaking conditions, providing support to music's ability to positively influence speech production (Ludke et al., 2014). If music is able to aid speech production of a novel language, it is plausible that music will also have an effect on speech perception, however continued research is necessary to better understand this relationship.

Receptive Language and Music

Music's effect on expressive language recovery has been the focus of much research, but little is known about the effects music may have on receptive language skills. To date, research on music and speech perception has primarily focused on how each set of stimuli is processed within the brain and whether they share neural computations or not. Very little research exists that behaviorally investigates whether music can aid an individual's ability to perceive speech. A study conducted by Schunk (1999) explored the effects of singing on receptive vocabulary skills for individuals learning English as a second language. Their results revealed that students acquired more target vocabulary words when the words were sung to them versus when the words were spoken. This research suggests that music has a positive effect on language acquisition but leaves many questions unanswered, including if this effect is seen in other populations and what aspects of music have the greatest effect on speech perception.

Other research has investigated the effects of musical training programs on the acquisition of receptive language and vocabulary in a number of populations. Bygrave (1995) conducted a study that examined the effects of a musical training program on the receptive language skills of students with reading difficulties. The experimental group participated in a 7-week intervention implemented by the classroom teacher that focused on students playing instruments, singing, and participating in creative movement and listening activities. The control group also participated in an intervention focused on story telling exercises that aimed to increase the development of early reading skills including listening, organization, comprehension, and memory (Bygrave, 1995). The musical intervention group showed more gains in receptive vocabulary skills than the story telling group. It is plausible to conclude that the musical training the experimental group received increased their speech segmentation skills as research conducted by Francois, Chobert, Besson, and Schön (2013) showed that musical training had a positive effect on the speech segmentation skills of children. Thus, music may play an important role in the development of receptive language skills by facilitating speech perception.

While research on children suggests that musical training increases a child's speech segmenting ability, no significant differences have been found in adults with and without musical training (Francois et al., 2013). However, musical training may have a positive effect on speech perception in adult musicians as there is evidence that musical training may facilitate enhanced auditory encoding skills in adult musicians (Kraus & Chandrasekaran, 2010). This speech perception improvement is achieved by increasing attention and executive functioning skills, which consequently decreases the interference effects of adjacent syllables/tones (Berti, Münzer, Schröger, & Pechmann, 2006). In line with this possibility, Berti et al., (2006) found that musicians were able to maintain a clearer tonal representation in the presence of tonal distractions compared to non-musician controls indicating stronger executive functioning and attention skills in musicians. Parbery-Clark, Skoe, Lam, and Kraus (2009) corroborated Berti et al.'s (2006) results when they found that musicians showed a heightened ability to comprehend speech in noisy environments compared to non-musician controls matched for hearing ability suggesting that increased attention and executive functioning skills in adult musicians aids their ability to comprehend language in demanding listening environments.

Priming studies provide some insight into this issue of whether music can aid receptive language in adults with damage to typical receptive speech networks. There are many different types of priming but the general premise is as follows: The initial stimulus ("the prime") activates a given set of brain networks. If the next stimulus presented ("the target") utilizes the same networks, the perception of the target will be facilitated (i.e. require less activation to respond to it) because the necessary brain networks are already firing; priming is not an all-or-none system, but rather facilitation is determined by the degree of relatedness between the prime and target (i.e. degree of overlap of brain networks). Numerous studies of both speech and music report a variety of priming effects. For example, Abad, Noguera, and Ortells (2003) reported that priming occurs when words are highly associated. Their two experiments examined the effects of priming as it relates to category and semantic association and found that priming effects are most profound when words are highly associated (Abad et al., 2003). Categorical relatedness was not sufficient in achieving strong priming effects; instead the high degree of association between the prime and target is the critical component (Abad et al., 2003). A review of the literature by Hutchinson (2003) corroborates these findings reporting that association is critical to automatic semantic priming and the strongest priming effects occur for words that are functionally related (i.e. synonyms and antonyms).

Priming is a common strategy used by researchers with lesion patients to investigate if they can take advantage of semantic, phonological, or grammatical contexts to help individuals with lesions access target words (Bates, Marangolo, & Pizzamiglio, 2001). A popular therapy technique, semantic feature analysis (SFA), is based on the use of semantic priming to help aphasics retrieve target words. Anomia is a salient characteristic of aphasia and it is believed to stem from broad based semantic deficits often found in aphasia (Chapey, 2010). SFA works by having the patient recall words semantically related to the target word until they are able to evoke the target word. The cueing technique is proscribed and patients are instructed to name words related to the association, group, action, properties, location, and use for the target word (Chapey, 2010; Davis & Stanton, 2005). The overall idea is that the semantically related words will prime the patient for retrieval of the target word.

The evidence for linguistic priming is strong and similar effects have been found in studies of harmonic priming (Tillman, Janata, & Bharucha, 2003; Justus & Bharucha, 2001). Behavioral studies using consonant and dissonant melodies demonstrate that related primes (harmonically similar chords) increase accuracy and decrease response time when compared to unrelated primes (Tillman et al., 2003). Using closely and distantly related chord primes and targets, Justus and Bharucha (2001) found that closely related primes and targets resulted in decreased response latencies when compared to distantly related primes and targets. They propose that primes create expectancy effects and that this expectancy effect allows for the decrease in processing speed when the prime and target are closely related (Justus & Bharucha, 2001). Comparable results have been found in studies of syntactic priming. Research by Smith and Wheeldon (2001) found that participants had decreased latencies when the prime and target sentence were syntactically related and increased latencies when they differed syntactically (Smith & Wheeldon, 2001).

It is logical to propose that just as semantic and syntactic priming occur within the domain of language, priming speech with music may be possible. For example, a musical melody primes the neural computations for processing the linguistic stimuli that follows. Steinbeis and Koelsch (2008) provide evidence that music can be a prime for language. They examined the N400 response elicited by musical chords and target words and report that the emotion expressed in music can prime the processing of language at the semantic level, as indicated by a smaller N400 response, which signifies closely related stimuli (Steinbeis & Koelsch, 2008). While music does not transmit the same level of semantic information as language, these EEG results indicate that music influences the processing of words and acts as a prime for language (Steinbeis & Koelsch, 2008). Results from this research corroborates earlier research (Koelsch et al., 2004): unrelated musical primes and target sentences elicit larger N400 responses compared to related musical primes and linguistic targets, implying more efficient processing of related primes and targets across domains. Additionally, the N400 response for incongruent primes and targets elicited the same increased response for both linguistic and melodic targets; this indicates semantic similarity and the ability of music to prime language when the melodic prime and lexical target are semantically related (Steinbeis & Koelsch, 2008). The degree of similarity between the two domains may be what drives the musical priming phenomenon and subsequently aids speech perception.

Present Study

Music and language are interconnected within the brain. The two domains likely engage shared and distinct neural resources. The commonalities between music and language within neural computations may contribute to the efficacy of MIT for expressive language recovery in non-fluent aphasia. If music is able to augment the recovery of expressive language, it may be able to do the same for receptive language. The notion of musical priming for language has provided support for music's ability to aid speech perception, but this phenomenon has rarely been looked at behaviorally, creating numerous questions about its functionality as a treatment intervention.

This study aims to investigate the effects of music on speech perception in demanding situations. In particular the study will investigate if music has a positive effect on speech perception and if so, what aspects of music contribute to improved speech perception. It is hypothesized that music will increase an individual's perception of spoken pseudoword sentences and that musical priming will have the greatest effect on speech perception. To answer these questions, participants listened to and overtly reproduced pseudoword sentences presented with and without rhythmic and melodic musical features following a 15 second time delay. The verbal repetition component of the present study is designed to drive activation of the dorsal stream of Hickok and

Poeppel's (2000) dorsal-ventral model, which is important in order to draw conclusions between the present study and MIT.

The present study includes both younger and older adult subjects. It is important to look at the effects of music on speech perception in the older adult population, as results in older adults will provide a more concrete understanding of how music may affect speech perception in individuals with aphasia: generalization between aphasics and older adults is easier as older adults exhibit deficits associated with healthy aging (Wingfield & Grossman, 2006; Grady & Craik, 2000) and many of these deficits are common in aphasia including reduced working memory and auditory comprehension skills (Potagas, Kasselimis, & Evdokimidis, 2011; Chapey, 2010). Furthermore, research on older adults is important because 75% of cerebral vascular accidents occur in individuals over the age of sixty-five (The Internet Stroke Center at Washington University, 2014), thus understanding how to facilitate speech perception is more valid when studied in the older adult population.

Methods

Participants

Sixty-one Arizona State University undergraduate students and 33 adults from the local community participated in the study. To be included in the study participants needed to be a) right-handed, b) a native speaker of American English, and c) have working memory within a normal range (defined below). Of the 94 participants who participated in the experiment, data from 20 participants was excluded from the final sample due to changes in the research design (*n=8*), exceptionally low working memory *(n=3)*, not following experimental procedures *(n=4)*, and ages outside the target ranges (*n=5)*, which left 74 participants with usable data. Participants ranged in age from 18-85 years (*M=48.58*) and included 69 females and 25 males. Participants included in the final sample were 18 males and 56 females ranging in age from 18-85 years (*M=46.76*). Participants received course extra credit or monetary compensation for their participation. Arizona State University's Institutional Review Board approved all procedures.

Participants were divided into two groups for analysis: younger adults and older adults. The younger adult group (n=50; 44 females) ranged in age from 18-37 years (*M=21.9)*. The older adult group (n=24; 12 female) ranged in age from 61-85 years *(M=71.6)*. Participants were non-musicians except for one participant in the older adult group who was a retired music teacher. The retired music teacher's data was included in the final sample as her data was not an outlier and did not quantitatively change the results.

Stimuli Design

Stimuli consisted of forty pseudoword sentences that consist of 5-9 words (*M= 6.3 words*; syllable length range = 7-13, *M= 9.7*). The pseudoword sentences were the same linguistic stimuli used in Rogalsky et al. (2011). The pseudoword sentences were created by replacing the content words of normal, syntactically correct sentences with phonologically plausible, yet meaningless words (i.e. Lantiff by mornicon will skinder.) (Rogalsky et al., 2011). Sentences were 2.55-4.31 seconds long (*M=3.26*). For each sentence a novel melody was composed by a classically trained vocalist and musician. Melodies were 3.54-4.90 seconds long (*M=4.38*).

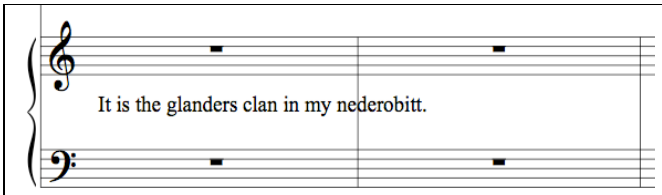
Four different conditions were created based on these sentences and melodies: sentences spoken with typical sentence prosody (normal speech), sentences spoken to the rhythm of the piano melody with simultaneous presentation of the melody (simultaneous music-speech), sentences spoken to the rhythm of the piano melody but no piano is heard (rhythmic speech), and presentation of the piano melody followed immediately by presentation of the sentence spoken to the rhythm of the melody (music-primed speech). Each condition was designed to include specific aspects of musical features except for the normal speech condition, which was intended to serve as a control. The rhythmic speech condition was designed to highlight the musical feature of rhythm. Rhythm was isolated as some research suggests that rhythm is the critical component of MIT (Stahl et al., 2011). The simultaneous music-speech condition was intended to highlight the musical features of rhythm and pitch in both domains with simultaneous presentation of the features and modes as other research indicates that both rhythm and pitch are critical to MIT (Zumbansen et al., 2014). Finally, the music-primed speech condition was intended to emphasize the musical features of rhythm and pitch in the two modes but with successive presentation of the modes as priming research implies that music is a viable prime for language (Steinbeis & Koelsch, 2008; Koelsch et al., 2004) in addition to pitch and rhythm being critical aspects of MIT (Zumbansen et al., 2014).

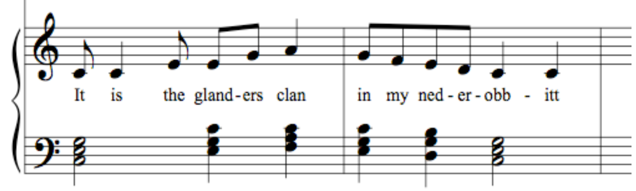
Melodies were two measures long and recorded at a tempo of 100 beats per minute. The melodies were single voice, followed traditional western chord progressions, and focused on the tonic in the key of C major. Piano accompaniments outlined the chord progression and followed the single voice melody line in the right hand. When spoken in rhythm, sentences followed the same rhythmic structure as its corresponding melody, which resulted in syllabic chunking of the linguistic component to the melodic rhythm. Stimuli were spoken and recorded by a female researcher using Audacity sound-editing software. Piano stimuli were generated using Finale SongWriter 2012. Audacity sound editing software was also used to combine recorded voice stimuli and piano stimuli.

Experimental Design

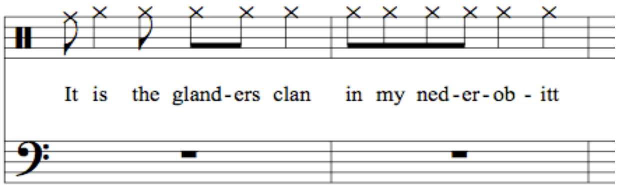
The experiment was designed using E-prime software 2.0 (Psychology Software Tools, Pittsburg, USA). All participants were binaurally presented with 40 stimuli, 10 of each type: normal speech, rhythmic speech, simultaneous music-speech, and music-primed speech. Four presentation orders were used to control for variability in pseudoword sentence duration and difficulty level. Each stimulus was presented one time and after a 15 second delay subjects were asked to repeat the sentence presented (Figure 1). The independent variable was the type of stimuli presented: normal speech, rhythmic speech, simultaneous music-speech, and music-primed speech. The dependent variable was the accuracy of the participant's verbal reproduction.

Participants were tested individually in a quiet room and were randomly assigned to one of the four presentation orders with 16-21 participants in each group. Participants were instructed to listen to each stimulus while viewing a fixation cross on the computer screen, and wait until the cross turned green before verbally reproducing the sentence. Participants were instructed not to sing or hum any of the accompanying melodies but otherwise were not provided with any instruction as to how to remember the sentence during the time delay.

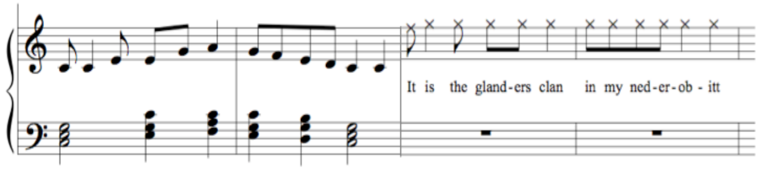
	15 second time delay	Verbal Reproduction

a) Normal Speech Condition: Sentences spoken with normal prosody and no music.

	15 second time delay	Verbal Reproduction

b) Simultaneous Music-Speech Condition: Sentences and piano melody presented simultaneously.

	15 second time delay	Verbal Reproduction

c) Rhythmic Speech Condition: Sentences spoken in rhythm to the composed piano melody, but no piano is heard.

	15 second time delay	Verbal Reproduction

d) Music-Primed Speech Condition: Presentation of piano melody immediately followed by spoken sentence. Note this condition is twice as long as the other conditions as the melody and sentence are presented successively.

Figure 1. Visual Representation of Conditions and Experimental Paradigm. Note condition descriptions under individual diagrams.

Working memory was tested using the Digit Span subtest of the *Wechsler Adult Intelligence Scale-IV (WAIS-IV)* (Wechsler, 2008) prior to the commencement of the E-prime experiment. Raw scores and scaled scores (*M=10; SD= +/- 3)* using *WAIS-IV* norms were computed from the Digit Span subtest, which asked participants to repeat numbers presented auditorally forwards, backwards, and in sequence. A cutoff normed score of ≥ 7 on the Digit Span subtest was used for inclusion in the final sample. All statistics presented in the results involving working memory performance used the raw Digit Span subtest scores.

All responses were video taped for later transcription. One researcher completed all the transcription and scoring. Reliability coding was conducted for 25% (*n= 18)* of the participants by a second researcher. The interrater reliability was 0.94. Responses were scored at the syllable level as either correct or incorrect using binary code. Incorrect responses were coded as any deviations from what was presented, including changes in vowels and plurals (i.e. tinder-tander, respendent-rependents) as well as real word and novel pseudoword substitutions. A proportion correct score was computed for each sentence (number of correct responses divided by total number of syllables). Proportion correct scores were calculated across the 10 stimuli in each condition to yield an overall proportion correct per stimulus type for each participant. Mean proportion correct scores were computed from all participants in each stimulus condition.

Results

Mean performances and standard errors for each stimulus type and age group are shown in Table 1 and Figure 2. It was hypothesized that musical priming would increase an individual's perception of the pseudoword sentences, while simultaneous presentation of musical elements would interfere with speech perception. A series of analyses of variance (ANOVA) and analyses of covariance (ANCOVA) were computed to test this prediction in both the younger adult and older adult groups. We then examined how working memory performance may be interacting with these results.

Table 1

Means and Standard Error of the Mean for Each Group

Stimulus	Younger Adults *(n=50)*	Older Adults *(n=24)*
Normal Speech	*M*=0.69, *SEM*=0.02	*M*=0.48, *SEM*=0.03
Rhythmic Speech	*M*=0.60, *SEM*=0.02	*M*=0.37, *SEM*=0.03
Simultaneous Music-Speech	*M*=0.61, *SEM*=0.02	*M*=0.35, *SEM*=0.03
Music-Primed Speech	*M*=0.65, *SEM*=0.02	*M*=0.47, *SEM*=0.03

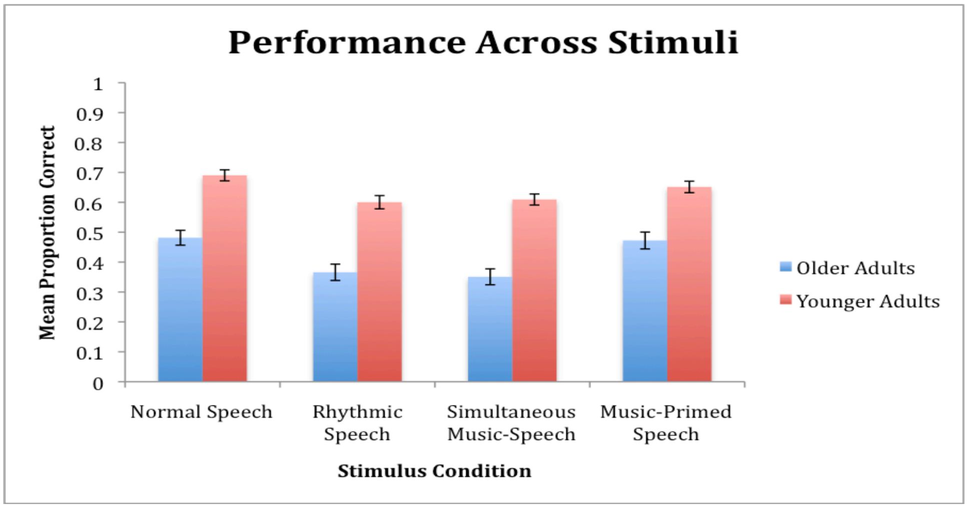


Figure 2. Mean proportion correct and standard error bars for each stimulus type.

Within-Age Group Findings

A repeated measures ANOVA was conducted in both the younger adult and older adult groups, respectively, to assess if stimulus type significantly influences participants' performance. In the younger adult group, Mauchly's test indicated sphericity had been violated, $\chi 2(5) = 13.30$, $p < .05$, therefore degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity ($\varepsilon = 0.83$). There was a significant main effect for stimulus type, $F(2.50, 122.60)=10.45$, $p<.001$, partial eta squared=.18. Post hoc tests using a Bonferroni correction for multiple comparisons revealed the significant contrasts (Table 2): normal speech performance in the young adults was significantly greater than rhythmic or simultaneous music-speech, and music-primed speech performance was significantly greater than the simultaneous music-speech condition.

In the older adult group, Mauchly's test indicated sphericity had not been violated, $\chi 2(5) = 2.99$, $p>.05$. There was a significant main effect for stimulus type, $F(3, 69)=12.21$, $p<.001$, partial eta squared=.35. Post hoc tests using the Bonferroni correction identified the significant contrasts (Table 3): normal speech performance in the older adults was significantly greater than rhythmic or simultaneous music-speech, and the music-primed speech condition was significantly greater than the simultaneous music-speech and rhythmic speech conditions.

Table 2

Younger Adult Group Bonferroni Correction Post Hoc Tests

Stimuli	*Mean Difference*	Standard Error	*p*
Normal Speech > Rhythmic Speech	.09	.02	.00*
Normal Speech > Simultaneous Music-Speech	.08	.02	.00*
Normal Speech > Music-Primed Speech	.04	.02	.19
Simultaneous Music-Speech > Rhythmic Speech	.01	.02	1.00
Music-Primed Speech > Simultaneous Music-Speech	.04	.02	.04*
Music-Primed Speech > Rhythmic Speech	.05	.02	.15

*Statistically significant.

A comparison of the three music-speech conditions suggests that musical priming facilitates greater performance than musical elements presented simultaneously with speech. Both the younger and older adult groups demonstrated a trend towards higher performance in the music-primed speech condition: the older adults' contrasts reached significance for music-primed speech being greater than both simultaneous music-speech and rhythmic speech (Table 3) while the younger adults' only reached significance for music-primed speech versus simultaneous music-speech (Table 2). Notably, the music-primed speech and normal speech conditions did not significantly differ in either group indicating that musical priming does not interfere with speech perception. These null results are remarkable because the musical priming condition contains rhythmic speech, but rhythmic speech alone yields significantly lower performance than the normal condition.

Table 3

Older Adult Group Bonferroni Correction Post Hoc Tests

Stimuli	*Mean Difference*	Standard Error	*p*
Normal Speech > Rhythmic Speech	.12	.03	.001*
Normal Speech > Simultaneous Music-Speech	.13	.02	.00*
Normal Speech > Music-Primed Speech	.01	.03	1.00
Rhythmic Speech > Simultaneous Music-Speech	.02	.02	1.00
Music-Primed Speech > Simultaneous Music-Speech	.12	.03	.002*
Music-Primed Speech > Rhythmic Speech	.11	.03	.01*

*Statistically significant.

Effects of Working Memory

Working memory was assessed for two reasons: (1) the neural correlates of working memory and sensory-motor integration have a high degree of overlap (Buchsbaum et al. 2011) and (2) it is well established that working memory performance declines in older adults (Grady & Craik, 2000). The younger adult group's working memory raw scores ranged between 22-40 (*M=27.72, sd=4.67)* (Figure 3). The older adult group had working memory raw scores ranging from 18-36 *(M=25.71, sd=4.67)* (Figure 3). An independent samples t-test of these working memory scores between the younger and older adult groups was significant, *t* (72)=1.72, *p* = .04 (one-tailed) indicating that as expected, the two age groups did have significantly different working memory capacities (Figure 4).

A Pearson product-moment correlation coefficient was computed to assess the relationship between performance in each condition and working memory in younger and older adults. In the younger adults, there was a significant ($p < .05$) positive correlation between performance in each condition and working memory: normal speech, *r*(48)=.50, *p*=.0002; rhythmic speech, *r*(48)=.36, *p*=.01; simultaneous music-speech, *r*(48)=.41, *p*=.003; music-primed speech, *r*(48)=.39, *p*=.005. In the older adult group, there also was a significant positive correlation between performance in each condition and working memory: normal speech, *r*(22)=.53, *p*=.008; rhythmic speech, *r*(22)=.64, *p*=.0008; simultaneous music-speech, *r*(22)=.45, *p*=.03; music-primed speech, *r*(22)=.42, *p*=.04. All correlations were positive in both age groups, indicating that higher working memory scores correlated with more accurate repetition in all conditions.

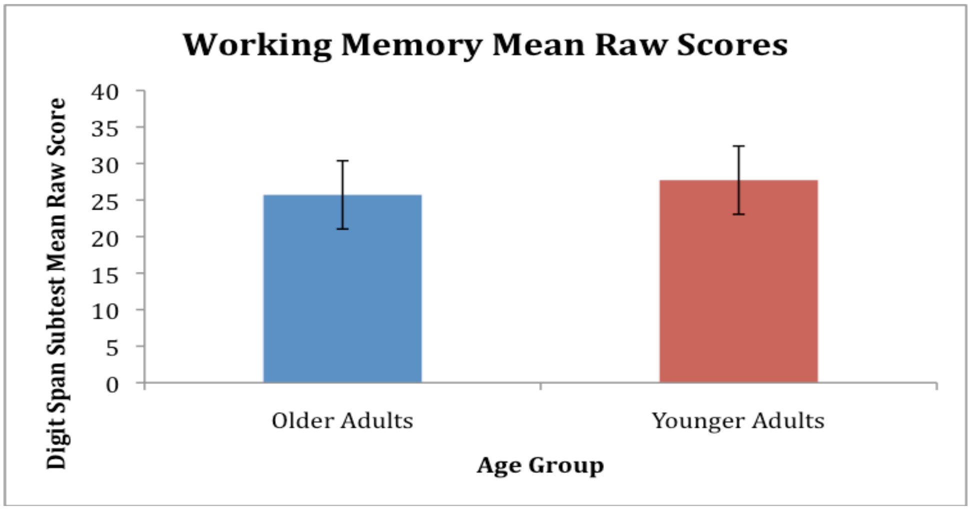


Figure 3. Working Memory Mean Raw Scores and Standard Deviations by age group.

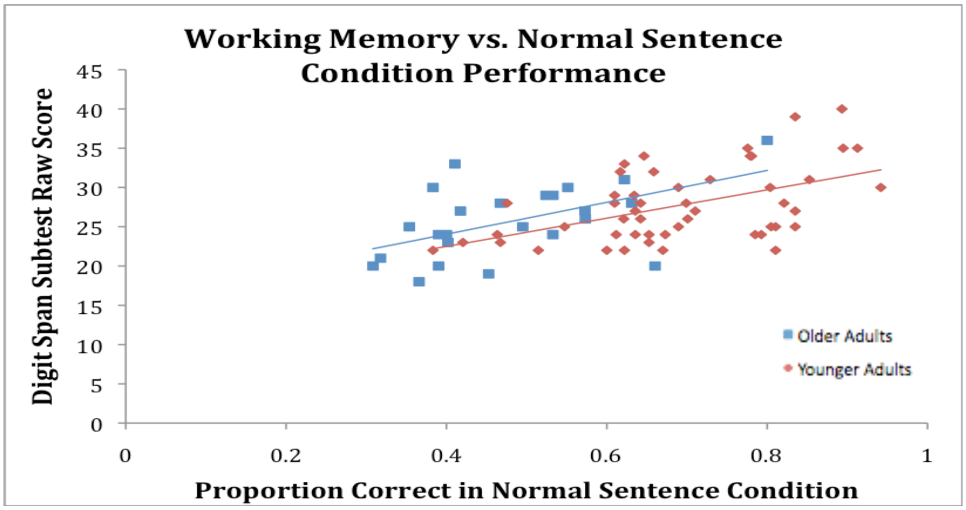


Figure 4. Working Memory Scatterplot for Younger and Older Adult Groups.

Age-Related Performance Differences

A mixed 2x4 ANOVA was conducted to assess how age group and stimulus type influenced performance. The within-subjects factor was stimulus type and the between subjects factor was age group. Mauchly's test indicated that the assumption of sphericity had not been violated, χ2(5) = 10.33, *p* > .05. There was a significant main effect for stimulus type, *F*(3, 216)=22.23, *p*<.005, partial eta squared=.24, corroborating earlier within-age group results. There was a significant between-subject main effect for age group, *F*(1, 72)=64.30, *p*<.005, partial eta squared=.47; pairwise comparisons indicate that younger adults performed better in each condition than the older adult group (Figure 2; Table 4). The interaction effect between stimulus condition and age group was not significant, *F*(3, 216)=2.20, *p*=.09, partial eta squared=.03.

A mixed 2x4 ANCOVA was conducted to assess how working memory may interact with age and stimulus type. The within-subjects factor was stimulus type, the between-subjects factor was group, and the covariate was working memory. Mauchly's test indicated that the assumption of sphericity had not been violated, $\chi 2(5) = 9.98$, $p > .05$. There was a significant main effect for group, *F*(1, 71)=67.54, *p*= .00, partial eta squared=.49, indicating that younger adults performed significantly better in each condition than the older adult group, even after removing working memory effects (Table 5). The interaction between age group and stimulus type was not significant, *F*(3, 213)=2.07, *p*=.11, partial eta squared=.03. The interaction between stimulus type and working memory was not significant, *F*(3, 213)= .30, *p*= .83, partial eta squared= .00.

Table 4

Younger Adult vs. Older Adult Group Bonferroni Correction Post Hoc Tests

Stimuli	*Mean Difference*	Standard Error	*p*
Normal Speech Younger > Normal Speech Older	.21	.03	.00*
Rhythmic Speech Younger > Rhythmic Speech Older	.23	.04	.00*
Simultaneous Music-Speech Younger > Simultaneous Music-Speech Older	.26	.03	.00*
Music-Primed Speech Younger > Music-Primed Speech Older	.18	.03	.00*

*Statistically significant.

Table 5

Younger Adult vs. Older Adult Group Bonferroni Correction Post Hoc Tests with Working Memory Covariate

Stimuli	*Mean Difference*	Standard Error	*p*
Normal Speech Younger > Normal Speech Older	.18	.03	.00*
Rhythmic Speech Younger > Rhythmic Speech Older	.21	.03	.00*
Simultaneous Music-Speech Younger > Simultaneous Music-Speech Older	.23	.03	.00*
Music-Primed Speech Younger > Music-Primed Speech Older	.16	.03	.00*

*Statistically significant.

Discussion

In the present study, music was hypothesized to increase an individual's ability to perceive speech. This hypothesis was tested using pseudoword sentences presented with (rhythmic speech, simultaneous music-speech, music-primed speech) and without (normal speech) musical components. Performances in each of the musical-speech conditions (rhythmic speech, simultaneous music-speech, music-primed speech) were compared to one another and to normal speech sentences. In both age groups the contrasts of normal speech > rhythmic speech and normal speech > simultaneous music-speech were statistically significant, suggesting that the addition of musical information interacts and therefore impairs speech perception in healthy adults. However, there was no significant difference between normal speech and music-primed speech. The lack of difference in performance between the normal speech and music-primed speech suggests that the musical priming of speech does not distract from speech perception, and could potentially facilitate it when speech networks are disrupted.

It was also hypothesized that priming speech with music (the music-primed speech condition) would improve an individual's speech perception when compared to simultaneously presenting speech and music or incorporating musical rhythm alone into the speech (simultaneous music-speech and rhythmic speech conditions). In younger adults, the musical priming did elicit greater performance than the simultaneous music-speech condition, but did not significantly differ from the rhythmic speech condition. In older adults, the musical priming elicited greater performance than both the rhythmic speech and the simultaneous music-speech conditions. These results suggest that musical priming facilitates speech perception compared to the simultaneous presentation of music and speech, and that older adults are particularly susceptible to this effect.

The current findings as they relate to musical priming, MIT, working memory, and the neurobiology of language are discussed in the subsequent sections.

Musical Priming

The statistical significance of the contrasts of rhythmic speech versus music-primed speech and simultaneous music-speech versus music-primed speech in the direction of music-primed speech, as well as the insignificance of the rhythmic speech versus simultaneous music-speech contrast provides support for the hypothesis that priming with music increases an individual's perception of spoken pseudoword sentences. But why might priming with music facilitate speech perception? As discussed in the introduction, Steinbeis and Koelsch (2008) demonstrated the ability of music to prime language using the N400 ERP. Using the platform of emotion, Steinbeis and Koelsch (2008) examined the N400 response elicited by musical chords and linguistic stimuli. Their research revealed that the emotion expressed in both music and language elicited similar N400 responses, indicating a deeper semantic relationship between the two domains. This semantic relationship between music and language indicates that music can prime language when the melodic prime and lexical target are highly related (Steinbeis & Koelsch, 2008). It is proposed that the priming scenario in the present experiment moderately facilitated speech perception, because the rhythm and melody of the piano were closely related to the rhythm and melody of the linguistic component thus facilitating speech perception of the linguistic stimulus. However, simultaneous presentation of the music and speech did not show a similar facilitation because the rhythm and pitch of the speech and music were competing for shared neural resources.

Musical priming for speech is effective because music and language have shared neural computations. This is not surprising given that both music and spoken language are complex, sound based systems that are initially processed in the auditory cortices bilaterally (Slevc, 2012). There is ample evidence to suggest that pitch and rhythm processes are driving the use of shared resources: research on individuals with amusia without aphasia reveals deficits in interpreting prosody, which contains both aspects of pitch and rhythm (Patel, 2013; Jiang et al., 2012); similarly, the clinical efficacy of MIT suggests that pitch and rhythm facilitate speech production in individuals with non-fluent aphasias (Zumbansen et al., 2014; Conklyn et al. 2012; Schlaug et al., 2008). This connection between speech and music in lesion studies is also evident in the relationship between a composer's native language and the melodies of their musical compositions. This association has been demonstrated by measuring the pitch interval variability, the aspect of intonation found in musical compositions, in the melodies of a French and English composer. Pitch interval variability is more consistent in French and more variable in English and this finding was consistent in the musical compositions analyzed: the French composer created compositions with less pitch interval variability while the English composer created compositions with greater pitch interval variability, thus musical scores reflected the composer's native language (Patel, 2005). Similar results were found when investigating the relationship between music and native language in terms of rhythm (Patel & Daniele, 2003). Again, using French and English composers, Patel and Daniele (2003) found that the rhythm in a composer's music reflects the rhythm of their native language. Similar findings have been identified in studies of culture and rhythm with non-musicians. Hannon (2009) demonstrated listener's ability to identify which language instrumental music belonged to based on the musical rhythm. The similarities Patel (2005), Patel and Daniele (2003), and Hannon (2009) found between the rhythm and pitch of native languages and musical compositions suggest that music may be used as a prime for language, but interferes if presented simultaneously.

Implications for MIT

Understanding the relationship between music and language is essential to optimizing and evaluating MIT. MIT is a clinical treatment for non-fluent aphasia that utilizes music to increase expressive language output (see introduction for a more detailed description of MIT). Schlaug et al. (2008) proposes four ways in which MIT may function: reduced rate of speech, syllable lengthening, syllable chunking, and left hand tapping. A reduced rate of speech is achieved through the intoning of words as singing allows articulation to occur at a reduced rate (Schlaug et al., 2008). Similarly, singing promotes syllable lengthening which aids in more fluent articulation and syllable chunking as singing promotes the production of shorter words and phrases naturally due to breath segmentation. Left hand tapping works to keep rhythm, and prime the right sensorimotor network for articulation making verbal expression more fluent (Schlaug et al., 2008).

Schlaug et al.'s (2008) possible mechanisms of MIT can be examined in the present experiment. The rhythmic speech condition employed the use of syllable chunking. The simultaneous music-speech condition utilized syllable chunking in addition to syllable lengthening, and simultaneous melodic rhythm. The music-primed speech condition, the only condition that appeared to facilitate speech perception, used melodic rhythm in the music prime and target sentence, as well as syllabic chunking in the target sentence. Thus, all three music-speech conditions utilize the strategy of syllable chunking, which arises from rhythm, indicating that musical rhythm interacts with rhythm processing in speech perception: simultaneous melodic and speech rhythm interfere with one another, while the presentation of a melodic rhythm primes speech rhythm processing, thereby enhances speech perception.

Syllable chunking has been shown to be a successful strategy in helping individuals with aphasia comprehend and remember information over extended periods of time (Romani, Galluzzi, Bureca, & Olson, 2011). Chunking is a successful strategy because it breaks longer phrases and sentences up into smaller components that are easier to perceive and remember. The rhythm of music provides a natural means in which syllables may be chunked. The music-primed speech condition may have had the highest performance of the three musical conditions because it employed the use of rhythm in two modes, melodic and linguistic rhythm. The melodic rhythm of the music-primed speech condition provided the participant with an initial exposure to the linguistic rhythm prior to the presentation of the pseudoword sentence. This rhythmic priming may provide individuals with an advantage in processing the subsequent sentence as they have pre-exposure to the rhythm that will determine the chunking of syllables making the processing of the subsequent sentence simpler.

Notably, it seems that rhythm alone does not facilitate speech perception or production, and the present findings regarding the rhythmic speech condition coincide with this literature (Zumbansen et al., 2014; Schlaug et al., 2008). Using three patients with aphasia, Zumbansen and colleagues (2014) examined the effects of MIT (intoning sentences and left hand tapping), compared to rhythmically spoken sentences and left hand tapping, and sentences spoken with normal prosody and no hand tapping. They concluded that generalization was greatest in the MIT condition, which contained components of musical pitch and rhythm. When rhythm was isolated, generalization was not as strong as when the elements of pitch and rhythm were both present. While the generalization effect of MIT was strongest in all three patients, the rhythmically spoken sentences showed weak generalization in two of the three patients. Generalization of Zumbansen et al.'s (2014) findings is limited due to small sample size, but the present study behaviorally corroborates their findings within speech perception. Participants in the present study performed more poorly on the rhythmically phrased sentences when compared to the music-primed sentences. Together, these studies propose that rhythm alone is not enough to facilitate speech production and perception and that a combination of musical elements including pitch and rhythm is necessary for increasing both production and perception of speech.

Hand tapping was not formally looked at in the present experiment, but an informal survey at the end of each participant's testing session indicates that many participants used some form of tapping to keep the rhythm of the melody and pseudoword sentence. In MIT, hand tapping acts as a metronome as well as a prime for the sensorimotor network (Schlaug et al., 2008). In the present experiment, hand/foot tapping could have aided the participant in maintaining the rhythm and syllable chunks, which may have contributed to the increase in speech perception seen in the music-primed speech condition.

Underlying Neurobiology

In the present study, the relatively poorer performance for simultaneous music-speech compared to music-primed speech points to speech and music having shared neural computations. The simultaneous presentation of the music and speech caused interference, while in the music-primed speech condition music activated the same networks as speech, thereby requiring less effort to maintain this activation for the subsequently presented speech.

Neuroimaging evidence exists that supports these behavioral results. Koelsch et al. (2002) found activation for musical stimuli bilaterally in the anterior-superior insular cortices, inferior frontolateral cortices including Broca's area, STG including Wernicke's area, planum polare, Heschl's gyrus, planum temporale, and the right superior third of the superior temporal sulcus in response to a melodic (rhythm and pitch) discrimination task. These areas are also known to activate in response to linguistic stimuli (Rogalsky et al., 2011). This overlap in activation patterns for music and linguistic stimuli supports the behavioral results of the present experiment suggesting an overlap in the neural computations used for processing speech and music, specifically overlap of the neural computations of pitch and rhythm.

Contributions of Working Memory

Working memory ability clearly influenced the present findings: in both the younger and older adult groups working memory was positively correlated with performance in all conditions, although effect sizes did vary. The positive nature of all the correlations indicates that higher working memory scores correlate with higher performance. This trend suggests that working memory effects should be taken into account when considering how music may impact speech perception and MIT outcomes. However, it is important to note that while working memory and task performance were overall positively correlated, the small effects sizes suggest that an individual's performance was not entirely determined by working memory abilities and that other unknown factors contribute to an individual's ability to capitalize on music as an aid for speech perception.

In younger adults, the strongest correlation (moderately positive) with working memory was found for the normal speech condition, with higher working memory scores being associated with higher performance. In this same group, the weakest correlation was in the rhythmic speech condition, which capitalized on the musical component of rhythm. This pattern may suggest that the rhythmic speech condition may benefit younger individuals with lower working memory capacities as the chunking of syllables may have reduced working memory demands.

Conversely, in the older adult group, the strongest (moderately positive) correlation with working memory was found for the rhythmic speech condition, which suggests that greater working memory capacity is needed to take advantage of the rhythmic component of music to aid speech perception in older adults. This correlation is opposite of what was found in the younger adult group, which suggests that age may play a role in how music and working memory processes interact with speech perception. Remarkably, the weakest correlation within the older adult group regarding working memory was with the music-primed speech condition, which coincides with previous work in several domains indicating that priming is not mediated through working memory (Baqués, Sáiz, & Bowers, 2004).

One musical component incorporated in the music-primed speech condition, syllable chunking is a common strategy used to treat working memory impairments associated with aphasia (Romani et al., 2011). In the music-primed speech condition, syllables were chunked to correspond with the preceding melodic rhythm. The chunking of the syllables to the melodic rhythm provided the individual with double exposure to the target rhythm (first melodic, then linguistic) and thus it is likely that the music prime facilitated speech perception via rhythmic processes. Individuals with lower working memory may benefit the most from musical assistance, while individuals with greater working memory abilities were able to perform equally well in all conditions without assistance. Although further research is needed in this area, these findings lay the foundation for more extensive work on how working memory affects the effectiveness of using music and rhythm to facilitate speech processing, such as in MIT.

Limitations and Future Directions

It is important to note the methodological limitations of this research. One limitation of the experiment was that the pseudoword sentences were solely presented auditorally and this is not an accurate representation of natural, live speech. When we perceive speech in a natural environment we are able to both hear and visually observe the speaker and this audiovisual integration is integral for accurate speech perception as both systems work together to comprehend the spoken message. An auditory only approach was used in this experiment so that speech perception could be looked at in the absence of compensatory strategies afforded by a visual component such as lip reading. While we solely focused on auditory speech perception additional studies are needed to understand how audiovisual integration interacts with musical priming and rhythmic processing more generally.

A second limitation of the experiment was that perception was measured via production. Measuring perception with oral production complicates the experiment because in addition to speech perception errors, errors could also have occurred in the auditory-motor integration processes necessary to generate the motor plan of the verbal reproduction. This method was chosen as it facilitated a natural context in that it is common for individuals to repeat sentences not accurately perceived for clarification in the course of a conversation. It is also reasonable to assume that the current subjects, healthy adults, do not have any difficulties producing the pseudoword sentences, as the pseudowords are phonologically plausible English words.

Another aspect of the limitation of using speech production as a measure of perception is that the neural computations involved in speech perception leading to repetition differ from those involved in speech comprehension (i.e. speech perception does not equal speech comprehension). Syllable discrimination or identification tasks, have been found to doubly dissociate from auditory comprehension tasks (i.e. auditory word-picture matching tasks; Baker, Blumstein, & Goodglass, 1981). This dissociation may reflect the fact that phonologically-driven tasks do not necessarily engage the same speech perception processes used in everyday speech processing that require semantic processing, i.e., they are not an ecologically valid task (Hickok & Poeppel, 2007). Findings from across studies also suggest that there is a dissociation in the underlying neuroanatomy supporting these two types of tasks: Dronkers, Wilkins, Van Valin, Redfern, and Jaeger (2004) found that impairments on an auditory word-picture matching task (which requires the perception of speech sounds) were significantly associated with damage to the left middle temporal gyrus, while Schwartz, Faseyitan, Kim, and Coslett (2012) found both auditory discrimination and non-word repetition tasks to implicate regions in left pre and post-central gyri, inferior frontal gyrus, inferior parietal cortex, and the superior temporal gyrus. This dissociation is reflected in Hickok and Poeppel's (2000) dorsal-ventral model of speech processing: sub-lexical phonological processing utilized in repetition activates the dorsal stream in the left inferior parietal and left inferior frontal lobes, while sensory processes interact with lexical and semantic representations in the bilateral temporal lobes (Hickok & Poeppel, 2000). The idea that the neural computations of speech perception and comprehension are distinct is important as speech perception abilities may be intact but a disruption in the neural computations of speech comprehension will still prevent the patient from comprehending the message.

A final limitation of the present experiment is that it only contained a behavioral component, which limits the understanding of how and where neural computations processed the stimuli. As previously discussed, speech perception and comprehension have distinct neural computations and the use of neuroimaging techniques would increase understanding of how each type of stimulus was anatomically processed. A neuroimaging component would also increase understanding of where the breakdown in processing occurs (i.e. in regions implicated in spectrotemporal analysis, sensorimotor integration, speech comprehension, articulation, etc.) as no participant reached the performance ceiling in any condition. Understanding how and where these stimuli are being processed will also aid understanding of the neural computations involved in the processing of speech and melodic stimuli and provide insight into how one domain may assist the other when damage is sustained.

As previously discussed, future research is needed in this area to better understand how music may facilitate receptive speech abilities. Future research should employ the use of audiovisual stimuli and both male and female speakers to mimic natural speech. A neuroimaging component should be employed to help understand the unique activation patterns of each stimulus and how these activation patterns relate to the known patterns of activation for linguistic and melodic stimuli. To increase the generalizability of the experiment beyond healthy controls, lesion patients with a variety of lesions patterns should be tested. Using participants with below average working memory would also help to increase the generalizability of the research to lesion patient populations, as it is likely that individuals with aphasia have reduced working memory capacities (Potagas et al., 2011; Chapey, 2010). While the present study solely used healthy controls, it lays the foundation to examine the relationship between music and speech perception with aphasia.

Conclusion

As the relationship between language and music is more deeply investigated, it has become more apparent that the two domains have many properties in common including rhythm, pitch, prosody, and emotion. These commonalities between language and music have lead to the creation of clinical treatments such as MIT that rely heavily on the similarities between the two domains. While much more is known about the effects of music on speech production, little research exists exploring the relationship between speech perception and music. Using a behavioral approach, this research explored the relationship between music and speech perception in younger and older adults, in an attempt to shed some light on what aspects of music may facilitate or interfere with speech perception. The hypothesis that musical priming would facilitate speech perception was supported; participants scored the highest in the music-primed speech condition compared to the rhythmic speech and simultaneous music-speech conditions. The shared rhythmic properties as well as the resulting syllabic chunking may drive the priming of speech perception via music. Older adults performed in a similar manner as younger adults although older adults scored significantly lower across all conditions, however, these age-related performance differences are accounted for by differences in working memory. This research provides the foundation for understanding how language and music interact within neural computations. Understanding this relationship will help us to create speech and language treatments that focus on facilitating performance in one domain through the use of neural computations in the spared domain.

References

Abad, M. J., Noguera, C., & Ortells, J. J. (2003). Influence of prime–target relationship on semantic priming effects from words in a lexical-decision task. *Acta Psychologica*, *113*(3), 283-295.

Abrams, D. A., Bhatara, A., Ryali, S., Balaban, E., Levitin, D. J., & Menon, V. (2011). Decoding temporal structure in music and speech relies on shared brain resources but elicits different fine-scale spatial patterns. *Cerebral Cortex, 21*(7)*,* 1507-1518.

Archibald, L. M., & Gathercole, S. E. (2006). Short-term and working memory in specific language impairment. *International Journal of Language & Communication Disorders*, *41*(6), 675-693.

Baddeley, A. (2003). Working memory and language: An overview. *Journal of Communication Disorders, 36*(3)*,* 189-208.

Baker, E., Blumstein, S. E., & Goodglass, H. (1981). Interaction between phonological and semantic factors in auditory comprehension. *Neuropsychologia*, *19*(1), 1-15.

Bates, E., Marangolo, P., Pizzamiglio, L., & Dick, F. (2001). Linguistic and nonlinguistic priming in aphasia. *Brain and Language, 76*(1)*,* 62-69.

Baqués, J., Sáiz, D., & Bowers, J. (2004). Effects of working memory load on long‐term word priming. *Memory*, *12*(3), 301-313.

Belin, P., Van Eeckhout, P., Zilbovicius, M, Remy, P., Francois, C., Guillaume, S., Chain, F., Rancurel, G., & Samson, Y. (1996). Recovery from nonfluent aphasia after melodic intonation therapy: A PET study. *Neurology, 47*(6)*,* 1504-1511.

Berti, S., Münzer, S., Schröger, E., & Pechmann, T. (2006). Different interference effects in musicians and a control group. *Experimental Psychology, 53*(2)*,* 111-116.

Berwick, R. C., Friederici, A. D., Chomsky, N., & Bolhuis, J. J. (2013). Evolution, brain, and the nature of language. *Trends in cognitive sciences*, *17*(2), 89-98.

Besson, M., & Schön, D. (2001). Comparison between language and music. *Annals of the New York Academy of Sciences, 930*(1)*,* 232-258.

Brandt, A., Gebrian, M., & Slevc, L. R. (2012). Music and early language acquisition. *Frontiers in Psychology, 3,* 1-17.

Brown, S. (2001). Are music and language homologues? *Annals of the New York Academy of Sciences, 930,* 372-374.

Buchsbaum, B. R., Baldo, J., Okada, K., Berman, K. F., Dronkers, N., D'Esposito, M., & Hickok, G. (2011). Conduction aphasia, sensory-motor integration, and phonological short-term memory–an aggregate analysis of lesion and fMRI data. *Brain and language*, *119*(3), 119-128.

Burunat, I., Alluri, V., Toiviainen, P., Numminen, J., & Brattico, E. (2014). Dynamics of brain activity underlying working memory for music in naturalistic condition. *Cortex, 57,* 254-269.

Bygrave, P. L., (1995). Development of receptive vocabulary skills through exposure to music. *Bulletin for the Council of for Research in Music Education, 127,* 28-34.

Carrington, J. F. (1971). The talking drums of africa. *Scientific American, 225*(6)*,* 90-94.

Chapey, R. (2010). Language intervention strategies in aphasia and related neurogenic communication disorders. 5th ed, MD: Lippincott Williams & Wilkins.

Conklyn, D., Novak, E., Boissy, A., Bethoux, F., & Chemali, K. (2012). The effects of modified melodic intonation therapy on non-fluent aphasia: A pilot study. *Journal of Speech, Language, and Hearing Research, 55*(5)*,* 1463-1471.

Dalla Bella, S., & Peretz, I. (1999). Music agnosias: selective impairments of music recognition after brain damage. *Journal of New Music Research*, *28*(3), 209-216.

Davis, L. A., & Stanton, S. T. (2005). Semantic feature analysis as a functional therapy tool. *Contemporary Issues in Communication Science and Disorders, 32,* 85-92.

Dediu, D., & Ladd, D. R. (2007). Linguistic tone is related to the population frequency of the adaptive haplogroups of two brain size genes, ASPM and Microcephalin. *Proceedings of the National Academy of Sciences of the United States of America, 104*(26)*,* 10944-10949.

Dronkers, N. F., Wilkins, D. P., Van Valin, R. D., Redfern, B. B., & Jaeger, J. J. (2004). Lesion analysis of the brain areas involved in language comprehension. *Cognition*, *92*(1), 145-177.

Ellis, N. C., & Sinclair, S. G. (1996). Working memory in the acquisition of vocabulary and syntax: Putting language in good order. *The Quarterly Journal of Experimental Psychology: Section A*, *49*(1), 234-250.

Falk, D. (2004). Prelinguistic evolution in early hominins: Whence motherese? *Behavioral and Brain Sciences, 27*(4)*,* 491-541.

Fedorenko, E., Behr, M. K., Kanwisher, N. (2011). Functional specificity for high-level linguistic processing in the human brain. *Proceedings of the National Academy of Sciences of the United States of America, 108*(39)*,* 16428-16433.

Fedorenko, E., McDermott, J., & Kanwisher, J. (2012). Sensitivity to musical structure in the human brain. *Journal of Neurophysiology, 108*(12)*,* 3289-3300.

Fedorenko, E., Patel, A., Casasanto, D., Winawer, J., & Gibson, E. (2009). Structural integration in language and music: Evidence for a shared system. *Memory & cognition*, *37*(1), 1-9.

Francois, C., Chobert, J., Besson, M., & Schön, D. (2013). Music training for the development of speech segmentation. *Cerebral Cortex, 23*(9)*,* 2038-2043.

Gazzaley, A., Cooney, J. W., Rissman, J., & D'Esposito, M. (2005). Top-down suppression deficit underlies working memory impairment in normal aging. *Nature neuroscience*, *8*(10), 1298-1300.

Grady, C. L., & Craik, F. I. (2000). Changes in memory processing with age. *Current opinion in neurobiology*, *10*(2), 224-231.

Griffiths, T. D., Rees, A., Witton, C., Cross, P. M., Shakir, R. A., & Green, G. G. (1997). Spatial and temporal auditory processing deficits following right hemisphere infarction. A psychophysical study. *Brain*, *120*(5), 785-794.

Hannon, E. E. (2009). Perceiving speech rhythm in music: Listeners classify instrumental songs according to language of origin. *Cognition, 111*(3)*,* 403-409.

Hausen, M., Torppa, R., Salmela, V. R., Vainio, M., & Sarkamo, T. (2013). Music and speech prosody: A common rhythm. *Frontiers in Psychology, 4,* 1-16.

Hickok, G., Buchsbaum, B., Humphries, C., & Muftuler, T. (2003). Auditory–motor interaction revealed by fMRI: speech, music, and working memory in area Spt. *Cognitive Neuroscience, Journal of*, *15*(5), 673-682.

Hickok, G., & Poeppel, D. (2000). Towards a functional neuroanatomy of speech perception. *Trends in cognitive sciences*, *4*(4), 131-138.

Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience*, *8*(5), 393-402.

Hutchison, K. A. (2003). Is semantic priming due to association strength or feature overlap? A microanalytic review. *Psychonomic Bulletin & Review*, *10*(4), 785-813.

Jiang, C., Hamm, J. P., Lim, V. K., Kirk, I. J., Chen, X., & Yang, Y. (2012). Amusia results in abnormal brain activity following inappropriate intonation during speech comprehension. *PLOS ONE, 7*(7)*,* e41411.

Just, M. A., Carpenter, P. A., Keller, T. A., Eddy, W. F., & Thulborn, K. R. (1996). Brain activation modulated by sentence comprehension. *Science*,*274*(5284), 114-116.

Justus, T. C., & Bharucha, J. J. (2001). Modularity in musical processing: the automaticity of harmonic priming. *Journal of Experimental Psychology: Human Perception and Performance*, *27*(4), 1000.

King, J., & Just, M. A. (1991). Individual differences in syntactic processing: The role of working memory. *Journal of memory and language*, *30*(5), 580-602.

Koelsch, S., Gunter, T. C., Cramon, D. Y. V., Zysset, S., Lohmann, G., & Friederici, A. D. (2002). Bach speaks: a cortical "language-network" serves the processing of music. *Neuroimage*, *17*(2), 956-966.

Koelsch, S., Kasper, E., Sammler, D., Schulze, K., Gunter, T., & Friederici, A. D. (2004). Music, language and meaning: brain signatures of semantic processing. *Nature neuroscience*, *7*(3), 302-307.

Koelsch, S., Schulze, K., Sammler, D., Fritz, T., Müller, K., & Gruber, O. (2009). Functional architecture of verbal and tonal working memory: an FMRI study. *Human brain mapping*, *30*(3), 859-873.

Koelsch, S., & Siebel, W. A. (2005). Towards a neural base of music perception. *TRENDS in Cognitive Sciences, 9*(12)*,* 578-584.

Kotilahti, K., Nissila, I., Nasi, T., Lipiainen, L., Noponen, T., Merilainen, P., Huotilainen, M., & Fellman, V. (2010). Hemodynamic responses to speech and music in newborn infants. *Human Brain Mapping, 31*(4)*,* 595-603.

Kraus, N., & Chandrasekaran, B. 2010. Music training for the development of auditory skills. *Nature Reviews Neuroscience, 11*(8), 599–605.

Lim, K., Kim, Y., Lee, H., Yoo, J., Hwang, J., Kim, J., & Kim, S. (2013). The therapeutic effect of neurologic music therapy and speech-language therapy in post-stroke aphasic patients. *Annals of Rehabilitation Medicine, 37*(4)*,* 556-562.

Loui, P., Alsop, D., & Schlaug, G. (2009). Tone deafness: A new disconnection syndrome? *The Journal of Neuroscience, 29*(33)*,* 10215-10220.

Ludke, K. M., Ferreira, F., & Overy, K. (2014). Singing can facilitate foreign language learning. *Memory and Cognition, 42*(1)*,* 41-52.

Makuuchi, M., Bahlmann, J., Anwander, A., & Friederici, A. D. (2009). Segregating the core computational faculty of human language from working memory. *Proceedings of the National Academy of Sciences*, *106*(20), 8362-8367.

Mandell, J., Schulze, K., & Schlaug, G. (2007). Congenital amusia: An auditory-motor feedback disorder? *Restorative Neurology and Neuroscience, 25*(3)*,* 323-334.

Masataka, N. (2007). Music, evolution and language. *Developmental Science, 10*(1)*,* 35-39.

McMullen, E., & Saffran, J. R. (2004). Music and language: A developmental comparison. *Music Perception, 21*(3)*,* 289-311.

Nation, K., Adams, J. W., Bowyer-Crane, C. A., & Snowling, M. J. (1999). Working memory deficits in poor comprehenders reflect underlying language impairments. *Journal of experimental child psychology*, *73*(2), 139-158.

Norton, A., Zipse, L., Marchina, S., & Schlaug, G. (2009). Melodic intonation therapy: Shared insights on how it is done and why it might help. *Annals of the New York Academy of Sciences, 11169*(1), 431-436.

Parbery-Clark, A., Skoe, E., Lam, C., & Kraus, N. (2009). Musician enhancement for speech-in-noise. *Ear Hear. 30*(6), 653–661.

Patel, A. D. (2005). The relationship of music to the melody of speech and to syntactic processing disorders in aphasia. *Annals of the New York Academy of Sciences, 1060*(1), 59-70.

Patel, A. D. (2012). “Language, music, and the brain: a resource-sharing framework,” in *Language and Music as Cognitive Systems*, Oxford, UK: Oxford University Press, 204–223.

Patel, A. D. (2013). Sharing and nonsharing of brain resources for language and music. In *Language, Music, and the Brain: A Mysterious Relationship* (1st ed., pp. 329-356). Cambridge, MA: MIT Press.

Patel, A. D., & Daniele, J.R. (2003). An empirical comparison of rhythm in language and music. *Cognition, 87*(1)*,* 35-45.

Patel, A. D., Peretz, I, Tramo, M., & Labreque R. (1998). Processing prosodic and musical patterns: A neuropsychological investigation. *Brain and Language, 61*(1)*,* 123-144.

Pearce, J. M. (2004). Selected observations on amusia. *European neurology*,*54*(3), 145-148.

Peretz, I. (1990). Processing of local and global musical information by unilateral brain-damaged patients. *Brain*, *113*(4), 1185-1205.

Peretz, I., Belleville, S., & Fontain, S. (1997). Dissociations between music and language functions after cerebral resection: A new case study of amusia without aphasia. *Canadian Journal of Experimental Psychology, 51*(4)*,* 354-368.

Peretz, I., & Hyde, K. L. (2003). What is specific to music processing? Insights from congenital amusia. *Trends in cognitive sciences*, *7*(8), 362-367.

Peretz, I., & Kolinsky, R. (1993). Boundaries of separability between melody and rhythm in music discrimination: A neuropsychological perspective. *The Quarterly Journal of Experimental Psychology*, *46*(2), 301-325.

Peretz, I., Kolinsky, R., Tramo, M., Labrecque, R., Hublet, C., Demeurisse, G., & Belleville, S. (1994). Functional dissociations following bilateral lesions of auditory cortex. *Brain*, *117*(6), 1283-1301.

Piccirilli, M., Sciarma, T., & Luzzi, S. (2000). Modularity of music: evidence from a case of pure amusia. *Journal of Neurology, Neurosurgery & Psychiatry*, *69*(4), 541-545.

Potagas, C., Kasselimis, D., & Evdokimidis, I. (2011). Short-term and working memory impairments in aphasia. *Neuropsychologia*, *49*(10), 2874-2878.

Price, C. J. (2010). The anatomy of language: a review of 100 fMRI studies published in 2009. *Annals of the New York Academy of Sciences*, *1191*(1), 62-88.

Price, C. J. (2012). A review and synthesis of the first 20years of PET and fMRI studies of heard speech, spoken language and reading. *Neuroimage*, *62*(2), 816-847.

Raz, N., Rodrigue, K. M., Head, D., Kennedy, K. M., & Acker, J. D. (1997). Selective aging of the human cerebral cortex observed in vivo: Differential vulnerability of the prefrontal gray matter. *Cerebral Cortex, 7*(3), 268 –282.

Rogalsky, C., & Hickok, G. (2011). The role of Broca's area in sentence comprehension. *Journal of Cognitive Neuroscience*, *23*(7), 1664-1680.

Rogalsky, C., Matchin, W., & Hickok, G. (2008). Broca's area, sentence comprehension, and working memory: an fMRI study. *Frontiers in human neuroscience*, *2*.

Rogalsky, C., Rong, F., Saberi, K., & Hickok, G. (2011). Functional anatomy of language and music perception: Temporal and structural factors investigated using functional magnetic resonance imaging. *The Journal of Neuroscience, 31*(10)*,* 3843-3852.

Romani, C., Galluzzi, C., Bureca, I., & Olson, A. (2011). Effects of syllable structure in aphasic errors: Implications for a new model of speech production. *Cognitive psychology*, *62*(2), 151-192.

Salame, P. & Baddeley, A. D. (1989). Effects of background music on phonological short-term memory. *The Quarterly Journal of Experimental Psychology, 41*(1)*,* 107–122.

Schlaug, G., Marchina, S., & Norton, A. (2008). From singing to speaking: Why singing may lead to recovery of expressive language in patient's with broca's aphasia. *Music Perception: An Interdisciplinary Journal, 25*(4)*,* 315-323.

Schlaug, G., Marchina, S., & Norton, A. (2009). Evidence for plasticity in white matter tracts of patients with chronic Broca's aphasia undergoing intense intonation-based speech therapy. *Annals of the New York Academy of Sciences, 1169*(1)*,* 385-394.

Schunk, H. A. (1999). The effects of signing paired with signing on receptive vocabulary skills of elementary ESL students. *Journal of Music Therapy, 36*(2)*,* 110-124.

Schulze, K., Zysset, S., Mueller, K., Friederici, A. D., & Koelsch, S. (2011). Neuroarchitecture of verbal and tonal working memory in nonmusicians and musicians. *Human Brain Mapping, 32*(5)*,* 771-783.

Schulze, K., & Koelsch, S. (2012). Working memory for speech and music. *Annals of the New York Academy of Sciences, 1252*(1)*,* 229-236.

Schwartz, M. F., Faseyitan, O., Kim, J., & Coslett, H. B. (2012). The dorsal stream contribution to phonological retrieval in object naming. *Brain*, *135*(12), 3799-3814.

Semal, C., Demany, L., Ueda, K., & Halle, P. A. (1996). Speech versus nonspeech in pitch memory. *Journal of Acoustic Society of America, 100*(2)*,* 1132-1140.

Slevc, L. R. (2012). Language and music: sound, structure, and meaning. *Wiley Interdisciplinary Reviews: Cognitive Science*, *3*(4), 483-492.

Stahl, B., Kotz, S. A., Henseler, I., Turner, R., & Geyer, S. (2011). Rhythm in disguise: Why signing may not hold the key to recovery from aphasia. *Brain: A Journal of Neurology, 134,* 3083-3093.

Steinbeis, N., & Koelsch, S. (2008). Comparing the processing of music and language meaning using EEG and FMRI provides evidence for similar and distinct neural representations. *PLoS One*, *3*(5), e2226.

The Internet Stroke Center at Washington University. (2014). *Stroke Statistics.* Retrieved from http://www.strokecenter.org/patients/about-stroke/stroke-statistics/

Tillman, B., Janata, P., & Bharucha, J. J. (2003). Activation of the inferior frontal cortex in musical priming. *Cognitive Brain Research, 16*(2)*,* 145-161.

Trehub, S. E. (2003). The developmental origins of musicality. *Nature Neuroscience, 6*(7)*,* 669-673.

Wallin, N.L., Merker, B., & Brown, S. (2000). The origins of music. Cambridge, MA: MIT Press.

Wechsler, D. (2008). *Wechsler Adult Intelligence Scale-Fourth Edition (WAIS-IV)* (4th ed.). San Antonio, TX: Pearson.

Wingfield, A., & Grossman, M. (2006). Language and the aging brain: Patterns of neural compensation revealed by functional brain imaging. *Journal of Neurophysiology, 96*(6)*,* 2830-2839.

Zumbansen, A., Peretz, I., & Hébert, S. (2014). The combination of rhythm and pitch can account for the beneficial effect of Melodic Intonation Therapy on connected speech improvements in Broca’s aphasia. *Frontiers in human neuroscience*, *8,* 592.

Zatorre, R. J., & Baum, S. R. (2012). Musical melody and speech intonation: Singing a different tune. *PLOS Biology, 10*(7)*,* 1-6.