Learning and Retention of Novel Words in Musicians and Non-Musicians:

The Impact of Enriched Auditory Experience

on Behavioral Performance and Electrophysiologic Measures

by

Elizabeth Stewart

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved August 2017 by the
Graduate Supervisory Committee:

Andrea Pittman, Chair
Barbara Cone
Yi Zhou

ARIZONA STATE UNIVERSITY

December 2017

ABSTRACT

Music training is associated with measurable physiologic changes in the auditory pathway. Benefits of music training have also been demonstrated in the areas of working memory, auditory attention, and speech perception in noise. The purpose of this study was to determine whether long-term auditory experience secondary to music training enhances the ability to detect, learn, and recall new words.

Participants consisted of 20 young adult musicians and 20 age-matched non-musicians. In addition to completing word recognition and non-word detection tasks, each participant learned 10 nonsense words in a rapid word-learning task. All tasks were completed in quiet and in multi-talker babble. Next-day retention of the learned words was examined in isolation and in context. Cortical auditory evoked responses to vowel stimuli were recorded to obtain latencies and amplitudes for the N1, P2, and P3a components. Performance was compared across groups and listening conditions. Correlations between the behavioral tasks and the cortical auditory evoked responses were also examined.

No differences were found between groups (musicians vs. non-musicians) on any of the behavioral tasks. Nor did the groups differ in cortical auditory evoked response latencies or amplitudes, with the exception of P2 latencies, which were significantly longer in musicians than in non-musicians. Performance was significantly poorer in babble than in quiet on word recognition and non-word detection, but not on word learning, learned-word retention, or learned-word detection. CAEP latencies collapsed across group were significantly longer and amplitudes were significantly smaller in babble than in quiet. P2 latencies in quiet were positively correlated with word recognition in quiet, while P3a latencies in babble were positively correlated with word recognition and learned-word detection in babble. No other significant correlations were observed between CAEPs and performance on behavioral tasks.

These results indicated that, for young normal-hearing adults, auditory experience resulting from long-term music training did not provide an advantage for learning new information in either favorable (quiet) or unfavorable (babble) listening conditions. Results of the present study suggest that the relationship between music training and the strength of cortical auditory evoked responses may be more complex or too weak to be observed in this population.

ACKNOWLEDGMENTS

I wish to express my gratitude to the following individuals who assisted in the preparation and completion of this study.

Dr. Andrea Pittman, my PhD advisor, who recognized my potential as an investigator, and who guided me and challenged me so that I could live up to it.

Dr. Barbara Cone, who fostered my interest in auditory evoked potentials, shared her expertise, and provided the training, support, and even the equipment that allowed me to explore a technique that was new to me.

Dr. Yi Zhou, who provided me with an education in neurophysiology that helped me in the completion of this project, and that will continue to be an asset throughout my career.

Dr. Gavin Bidelman, who shared with me the audio files used in the evoked potential portion of this study.

Elizabeth Rainy, the research coordinator for the Pediatric Amplification Lab, who helped with participant recruitment.

Finally, I would like to acknowledge the study participants for their time and cooperation.

TABLE OF CONTENTS

Page

CHAPTER

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

STATEMENT OF THE PROBLEM

An extraordinary consequence of collective achievements in medical science is the advancement of human longevity. Extended life expectancy provides potential advantages not only for individuals, but also for the cultures in which they participate. Longer life spans have also, however, brought about new challenges related to health. One of the most common concerns, particularly in developed parts of the world, is the preservation of cognitive abilities. The rise in popularity of “brain-training” programs and interactive applications reflect a prevailing desire to maintain optimal cognition in older age. The role of physical activity, nutrition, and socialization in sustaining healthy cognitive function has also been an area of increased focus in recent years. Challenging the healthy brain in various ways appears to be a key to protecting against or minimizing the effects of future cognitive insult. Cognitive insult – a disruption in normal cognitive function – can result from a traumatic event, such as a brain injury or stroke, or may occur more organically, as in the case of aging or sensory deprivation secondary to hearing loss.

Age. Certain cognitive and perceptual abilities are known to decline with age. Speech comprehension, for example, appears to be less accurate in older compared to younger adults, even in listeners for whom hearing loss is not a factor. In a study of younger (~25 years old) and older (~70 years old) adults with clinically normal hearing, the older group performed more poorly on a behavioral measure of vowel discrimination using a five-step vowel continuum from /u/ to /a/ (Bidelman, Villafuerte, Moreno, &

Alain, 2014a). Thus, older adults' difficulty in discriminating amongst speech sounds may stem from changes in auditory processing and rather than reduced audibility.

Working memory has also been shown to decline with age. Jonides and colleagues (2008) identified three essential processes involved in working memory: encoding, maintenance, and retrieval. Encoding entails focusing attention on the incoming perceptual information and converting it to a cognitive representation. Maintenance stores this representation and prevents it from decaying, while retrieval involves bringing it back into cognitive focus (Jonides et al., 2008). Bialystok, Craik, and Luk (2008) tested the working memory of younger (~20 years old) and older (~68 years old) adults and found that older adults performed significantly more poorly than the younger group on two of the three tests used in the study (Bailystok et al., 2008). Similarly, McEvoy, Pellouchoud, Smith, and Gevins (2001) assessed working memory ability of young (~21 years old), middle-aged (~47 years old), and older (~69 years old) adults using a spatial working memory task. This measure is thought to reflect the speed at which the meaning of a stimulus can be integrated with an internal construct. Accuracy was notably (though not significantly) lower in older compared to younger adults, while reaction times were significantly longer in older adults than in the younger group, suggesting that central processing speed decreases with increasing age (McEvoy et al., 2001).

The effects of age on cognitive and perceptual abilities have been investigated in adults with music experience, to assess the potential of musical training to mitigate age-related decline in speech perception and working memory. For instance, Zendel and Alain (2012) assessed speech perception in varying levels of noise in musicians (aged 19-

91 years) and non-musicians (aged 18-86 years). Results revealed that, as age increased, musicians required less of an increase in SNR to maintain 50% performance compared to their age-matched non-musicians peers. Thus, musicians' ability to perceive speech in noise appears to decline less rapidly with age (Zendel & Alain, 2012).

Compared with speech perception advantages, the potential of music training to counteract diminishing working memory abilities is somewhat less straight-forward. For instance, Bugos, Perlstein, McCrae, Brophy, and Bedenbaugh (2007) examined the effect of individualized piano instruction (IPI) on working memory performance. Older adults with less than five years of musical training were recruited from an independent-living residential community and assigned to one of two groups: (1) an experimental group, receiving 6 months of IPI, consisting of a 30-minute lesson and at least 3 hours of independent practice each week, or (2) a control group, receiving no intervention. Four subtests from the third edition of the Wechsler Adult Intelligence Scale (WAIS-III; Wechsler, 1997) were used to assess working memory, and were administered to all participants prior to intervention and after the 6-month training period. Baseline scores did not differ across groups. On post-training assessment, participants in the experimental group showed an improvement over baseline performance on one of these subtests (Digit Symbols, which assesses the speed at which subjects are able to match the appropriate symbol to a list of digits, based on digit-symbol pairs presented in a table). These findings suggest that, even if initiated later in life, musical training has the potential to improve working memory capacity (Bugos et al., 2007). On the other hand, Hanna-Pladdy and MacKay (2011) used the Letter-Number Sequencing (LNS) subtest of the WAIS-III (Wechsler, 1997) to assess auditory working memory in older healthy adults. Subjects were divided into three groups: (1) non-musicians (no formal musical training), (2) low-activity musicians (1-9 years of musical experience, with some training), and (3) high-activity musicians (10+ years of musical training). In the LNS task, subjects were read a mix of numbers and letters and were asked to list the numbers in ascending order and the letters in alphabetical order. Contrary to the results of the Bugos et al. (2007) study, no significant differences between groups were observed on this measure. However, performance of the low-activity musicians fell between that of non-musicians and high-activity musicians, indicating a linear relationship between cognitive functioning in later adulthood and years of music experience (Hanna-Pladdy & MacKay, 2011).

Deficits in speech perception and working memory in older adults may be due to age-related changes in auditory processing. Using electrophysiologic measures, Pinal, Zurrón, and Díaz (2015) found effects of age on the encoding stage of working memory, as evidenced by smaller P3 responses to visual stimuli in older compared to younger adults. These results suggest that the processing resources needed to accurately categorize stimuli to be encoded into working memory may become depleted with advancing age.

It has been suggested that declining speech perception may be due to a decrease in neural inhibition (Bidelman et al., 2014a; Parbery-Clark, Anderson, Hittner, & Kraus, 2012), which has been shown to occur with advancing age (see Caspary, Ling, Turner, & Hughes, 2008 for review). Effectively, this decrease in inhibition increases temporal jitter (i.e., the variation in interspike intervals of the phase-locked neural response) in auditory processing, leading to distortion of the speech signal (Bidelman et al., 2014a;

Parbery-Clark et al., 2012). Likewise, Bidelman and colleagues found that a group of older normal-hearing adults, who showed poorer vowel discrimination than normal-hearing young adults, also displayed less robust auditory brainstem responses to the same stimuli compared to the younger group (Bidelman et al., 2014a).

In an examination of the potential physiological benefits of musical training, Bidelman and Alain (2015) recruited older adult (~70 years old) musicians and non-musicians to complete the same behavioral and electrophysiologic measures of speech discrimination used in the Bidelman et al. (2014a) study described above (i.e., 5-step vowel continuum from /u/ to /a/, with concurrent event-related potential recording). Groups did not differ in behavioral performance on vowel discrimination, though cortical responses revealed larger P3 amplitudes in musicians compared to non-musicians, suggesting that musical experience may enhance neural discrimination of speech sounds in older adults. These results prompted the authors to suggest that musical training may help to prevent age-related decreases in neural inhibition thought to be responsible for declining speech perception ability (Bidelman & Alain, 2015).

Hearing loss. The consequence of hearing loss is perhaps most clearly observed in deficits in speech perception abilities. Helfer and Wilber (1990) found both age and hearing loss severity to be negatively correlated with speech recognition in a group of adults over 60 years of age, despite finding no significant correlation between participants' ages and the degree of their hearing loss (i.e., pure tone average). This suggests that hearing loss effects speech intelligibility independently of age-related declines in auditory processing.

It is not just perceptual processing that suffers as a consequence of hearing loss. The pragmatics of communication – generally considered a higher-level cognitive skill – are also degraded by permanent hearing loss. Though this relationship may appear less direct than with speech perception, it is no less problematic. For example, older listeners with hearing loss identify emotions conveyed in speech less accurately than older listeners with normal hearing – a deficit which hearing aids do little to correct (Goy, Pichora-Fuller, Singh, & Russo, 2016).

Finally, uncorrected hearing loss impairs a listener's ability to learn new information – a skill that relies on both bottom-up and top-down processing, as described in the following section. Hearing loss slows the rate of language learning in children, resulting in a progressively widening gap in their receptive and expressive language abilities compared to their normal hearing peers (Blamey et al., 2001). Further, hearing loss has been shown to significantly decrease the ability to rapidly acquire new words in both adults (Pittman, Stewart, Willman, & Odgear, 2017) and children (Pittman, 2011; Pittman, 2008). With the use of amplification however, the number of exposures needed to learn new words decreased significantly in both children and adults, suggesting that prosthetic devices can compensate for the effect of peripheral hearing loss on learning to a certain degree.

In adults, it is often difficult to isolate the effects of hearing loss on perceptual and cognitive functions like speech perception in noise and working memory from those of normal aging, as prevalence and degree of hearing loss both increase in tandem with age. As a result, investigations of older hearing-impaired musicians and non-musicians have been fairly limited, and few if any studies have examined speech perception or related measures in this population. Such studies may be forthcoming, however, as recent work has examined the impact of hearing loss and music training on music-specific abilities. Moreno-Gómez and colleagues (2017) administered the Montreal Battery for the Evaluation of Amusia (MBEA; Peretz, Champod, & Hyde, 2003) to adults >60 years old with (*n*=42) and without (*n*=49) presbycusis. Eleven participants in each group (presbycusis, control) had at least one year of music training experience (presbycusis: mean=4.00 ± 5.50 years; control: mean=5.72 ± 6.33 years). The MBEA is a test of music processing, consisting of a series of subtests that assess listeners' ability to recognize and discriminate melodic (pitch scale, pitch contour, pitch interval) and temporal (rhythm, meter) dimensions of music. MBEA global (average) scores were significantly higher in the control group compared to the presbycusis group, suggesting that presbycusis does more to degrade music perception skills than aging alone. Additionally, the magnitude of impairment for melodic (pitch) dimensions of music in particular was smaller in musically-trained individuals with presbycusis compared to those with no history of music training (Moreno-Gómez et al., 2017).

Taken together, these studies suggest that music training may reduce the effects of age-related hearing loss, particularly with respect to the perception of familiar speech as well as the detection and learning of unfamiliar words. That is, music training and the challenges it imposes on sensory and cognitive processing may enhance skills important for learning new information. If such enhancements are established in earlier adult years, they may serve to protect against declining perceptual and cognitive abilities due to aging or hearing loss. The purpose of this experiment was to determine whether long-term auditory experience secondary to music training enhances the ability to detect, learn, and recall new words in young adults with normal hearing and cognitive function.

CHAPTER 2

THEORETICAL SUPPORT AND REVIEW OF THE LITERATURE

Word Learning

It has been suggested that auditory learning begins in infancy, and that prenatal auditory experience (exposure to structured speech stimuli) strengthens neural activity important for speech perception and language acquisition (Partanen et al., 2013). That is, the foundations of receptive language are already in the early stages of development at the time of birth.

Expressive vocabulary begins to develop when children are about 12 months old and continues throughout childhood and adolescence, yielding an average vocabulary of approximately 60,000 words by the time they graduate from high school (e.g., Pinker, 1994). The rate of new-word acquisition is slow in early childhood, accelerating over time and reaching a peak learning rate sometime between 10 and 17 years of age (Bloom, 2000). Word learning is a complex process dependent upon the interplay among conceptual, social, and linguistic abilities (Bloom, 2000). Thus, a singular theory has yet to emerge regarding the precise method by which children are able to go from 0 to 60,000 words in an 18-year period.

When has a word been learned? Word learning implies having an understanding of the meaning of a new word. According to Bloom (2000), knowing the meaning of a word requires: (1) having a mental representation or concept that the word symbolizes, and (2) mapping that concept onto the correct linguistic form (i.e., unit of speech capable of carrying meaning, such as a morpheme, word, or phrase). A theory put forth by Sternberg and Powell (1983) suggests that there are three components involved in novel word learning: selective encoding, selective combination, and selective comparison. Information relevant to learning a new word is rarely presented in isolation; rather, it is often embedded in a background of information, not all of which is necessary for acquiring the meaning of the novel word. Selective encoding is the process of examining all of the available information and identifying that which is critical for understanding the word's meaning, while filtering out irrelevant data. Selective combination involves taking the relevant pieces of information (obtained from selective encoding) and combining them into a cohesive unit. Finally, selective comparison describes the process of evaluating the new information to determine how it relates to existing knowledge (Sternberg & Powell, 1983). Essentially, novel information is selected, integrated, and examined to arrive at word meaning. But what are the sources of this new information?

Learning words from direct instruction. The most observable type of word learning occurs though deliberate teaching. In early childhood, this takes the form of ostensive naming – that is, labeling targets in the child's environment (e.g., "that's your *shoe*," "*throw* the ball"). A limitation of this approach is the problem of generalization – that is, the child must have a way to take a word that was learned in one specific situation and apply it appropriately in a new situation. In a classic example (Quine, 1960), a linguist is visiting a culture whose language has nothing in common with his own. A native of the culture points to a white rabbit that is scurrying by and says, "gavagai." With no common language between them, the linguist is unable to determine with certainty if "gavagai" means "rabbit," or "white," or "scurrying," or "animal," etc. (Bloom, 2000). A young child learning words faces the same dilemma when the parent directs their attention to the family dog and says, “dog.” Nothing in the act of simply naming the target makes it clear that “dog” refers to the dog, if the dog is also brown, soft, big, standing up, sniffing the ground, etc. In other words, labeling does not provide any clues as to how the word “dog” generalizes to novel situations.

Another method of deliberate instruction leading to word learning, typically beginning in about the fourth grade, is the memorization of vocabulary lists. This approach requires substantial effort on the part of both the learner and the instructor, as it usually involves careful explanation of definitions and repeated practice. Thus, learning words well enough to recognize them easily and use them correctly takes a considerable amount of time using this method. In fact, it is estimated that children learn only 100-200 words this way over the course of a school year (Miller & Gildea, 1987). From where, then, is the majority of a high-school graduate’s 60,000-word vocabulary learned?

Learning words from context. Most words are learned from linguistic context (Sternberg, 1987; Bloom, 2000). Sternberg and Powell (1983) proposed a model of contextual learning comprised of three components: (1) knowledge-acquisition processes, (2) contextual cues, and (3) moderating variables. Knowledge-acquisition processes consist of selective encoding, selective combination, and selective comparison (described above). These processes operate on a set of cues that are drawn from the context in which the new words are embedded.

Sternberg and Powell (1983) identify eight categories of contextual cues that can help to establish the meaning of an unknown word (“X”): (1) temporal cues, indicating when or how often X can occur, or the duration of X; (2) spatial cues, which inform the learner about the location of X (general or specific); (3) value cues, regarding the worth of X or the kinds of effect(s) X brings about; (4) stative descriptive cues, which provide information about X's physical properties (size, shape, etc.); (5) functional descriptive cues, alluding to potential purposes of or uses for X; (6) causal/enablement cues, identifying the things that can cause or enable X; (7) class membership cues, which indicate which class(es) X belongs to, or which identify other members of the class(es) to which X belongs; and (8) equivalence cues, which are cues regarding the meaning of X or contrasts to the meaning of X.

Moderating variables dictate the degree to which these cues facilitate word learning from context (Sternberg and Powell, 1983). One such variable is the number of times the unknown word is encountered. If the unknown word occurs multiple times, each time surrounded by different cues, information provided by those cues can be integrated to provide a clearer idea about what the unknown word means. As another example, different types of information about the unknown word can stem from the variability of contexts in which the word appears over multiple occurrences. Contexts that vary on factors such as subject matter or writing style are likely to provide a mixture of cues carrying unique information about the word's meaning. A third variable that influences contextual learning is how critical the word is to the understanding of the phrase, sentence, or passage in which it appears. If knowing the meaning of the unknown word is necessary to understand its surrounding content, the learner will be highly motivated to determine the word's meaning. Additionally, the relationship between the specific contextual cue and the nature of the unknown word determines, in part, how helpful that cue will be in deciphering the word's meaning. For example, a spatial cue will be more helpful than a temporal cue in reaching the conclusion that the unknown word refers to a particular location. Another moderating variable is the density of unknown words. If the learner is confronted with many unknown words within a single context, they will likely have considerable difficulty making correct pairings between contextual cues and the unknown words to which they apply. Finally, the usefulness of contextual cues is dependent upon the learner's ability to retrieve and apply prior knowledge in relating the cue to the unknown word (Sternberg and Powell, 1983). Thus, under this model, learning new words from context involves cognitive processing of multiple types of cues, which is moderated by variables involving the unknown word itself, its relationship with its context, and even the capabilities of the learner.

Sternberg and Powell (1983) tested this theory by having high school students read short passages, each containing 1 to 4 words with extremely low-frequency of occurrence in everyday speech. Subjects were then asked to provide a definition for each low-frequency word. Responses were given a "goodness-of-definition" rating. These ratings were then examined as a function of the number and strength of contextual cues and moderating variables that were available to help determine the meaning of the low-frequency word. As a whole, contextual cues and moderating variables were fairly strong predictors of goodness-of-definition scores (r=0.77-0.93; Sternberg & Powell, 1983).

Syntactic vs. non-syntactic context. Contextual cues may be classified on the basis of whether they are syntactic or non-syntactic in nature.

Non-syntactic context. Word learning from non-syntactic context occurs due to a sensitivity to the overall meaning of the phrase, sentence, or passage in which the novel word is embedded (Bloom, 2000). For example, in the sentence, "She is feeding a dax," the available context leads to the conclusion that a dax is a living thing rather than an inanimate object, as it is capable of being fed (Sloutsky & Yao, 2008). Beals (1997) examined the conversations of preschoolers' families (specifically, at mealtimes), paying particular attention to exchanges involving words with which the child was unlikely to be familiar. Results revealed a significant positive correlation ($p<0.01$) between the frequency of semantic support for rare words during mealtime discourse and the children's scores on the Peabody Picture Vocabulary Test (Dunn & Dunn, 1981) 2-3 years later. These results suggest that children are able to learn at least some words from spoken context (Beals, 1997). A study by Nagy, Herman, and Anderson (1985) aimed to measure word learning from non-syntactic context presented in a written format. Eighth-grade students read either an expository or a narrative text and were then administered two forms of vocabulary tests on target words contained in each of the texts (both the one they read and the one they didn't read). These tests consisted of an individual interview and a multiple-choice test, both designed to assess partial knowledge of word meaning. The degree to which students performed better on words from the read passage than from the unread passage was interpreted as the extent to which they were able to learn word meanings from written context. For both assessments (interview and multiple choice), subjects scored significantly higher ($p<0.01$) on words from read passages than from unread passages. In other words, the students were able to learn at least partial meanings of words from non-syntactic context alone (Nagy et al., 1985).

Syntactic context. Word learning via syntactic context involves the use of cues provided by the grammatical structure of the phrase or sentence in which an unfamiliar word appears. That is, the word's function within the sentence can provide information about the word's meaning. Typically, when an unfamiliar word is presented, its context makes it clear in which part of speech the unknown word belongs (Brown, 1957). In the example given above, it is easy to determine that the novel word ("dax") is a noun. Syntactic cues act as a filter that directs the learner's attention towards features that are relevant to the new word's meaning (Brown, 1957). In a study of 16 pre-school aged children, Brown (1957) examined how assignment of words to particular parts of speech impacts word learning. Children in this study were first shown a picture depicting an action (verb), a substance (mass noun), and an object (particular noun) – for example, a pair of hands kneading (verb) a container (particular noun) filled with confetti (mass noun). Children were then taught a new (nonsense) word for one of the picture elements. For example, the word "sib" was used to describe the particular noun, mass noun, or verb. When "sib" was used as the verb (i.e., to refer to kneading), the examiner would ask the child, "Do you know what it means to sib?" The examiner would then show the child the picture and say, "In this picture, you can see sibbing." The particular noun and mass noun were introduced as "a sib" and "some sib," respectively. The children were then shown three additional pictures, each depicting only one of the elements in the composite picture and were asked to select the picture that showed what had been named in the first picture (e.g., if "sib" was presented as a verb, the child was asked to "show me another picture of sibbing"). Ten (63%) of the 16 children correctly selected the image showing the action when the word was presented as a verb, 11 children (69%) correctly selected the object when the word was presented as a particular noun, and 12 children (75%) correctly selected the substance when the word was presented as a mass noun. These results suggest that syntax can act as a cue for the general meaning of a novel word. Certainly, knowing that an unfamiliar word is a noun rather than a verb is not itself sufficient to determine the word’s precise meaning (Brown, 1957). Yet the ability to identify the part of speech that characterizes the novel word may be necessary for successful word learning (Bloom, 2000).

Fast-mapping vs. slow-mapping. Closely related to contextual learning is the concept of fast-mapping, in which children learn the meaning (or part of the meaning) of a new word after only one or two exposures, and without formal instruction (Carey & Bartlett, 1978; Heibeck & Markman, 1987). Fast-mapping has been offered as a partial explanation for the rapid increase in vocabulary during childhood (e.g., Rice, Buhr, & Nemeth, 1990). There is some question, however, as to whether children’s ability to link novel words with the concept (object, event, etc.) to which they refer (i.e., the referent) truly constitutes learning (Horst & Samuelson, 2008). To answer this question, Horst and Samuelson (2008) tested novel word retention in 2-year-old children. Children were presented with sets of three objects: one named familiar object, one unnamed familiar object, and one named novel object. For each set, children were asked to “get the [novel object]” (referent task). After a 5-minute delay, children were presented with sets containing two previously-seen novel objects and one previously-seen unnamed familiar object and were asked to “get the [one of the novel objects]” (retention task). Children’s performance on the referent task was significantly higher than chance ($p<0.0001$), but performance on the retention task was not, suggesting that fast-mapping does not produce a sufficiently salient representation of word meaning to allow it to be retained over a 5-minute period (Horst & Samuelson, 2008). It is argued, therefore, that both fast- and slow-mapping are needed for word learning (e.g., Axelsson, Churchley, & Horst, 2012). Slow-mapping occurs when the semantic representation of the novel word develops over time as a result of multiple and different kinds of experiences with the word (Capone & McGregor, 2005). While fast-mapping is important for establishing an association between a novel word and its referent, slow-mapping allows the full meaning of the word to be learned (Axelsson et al., 2005; Capone & McGregor, 2005).

Music Learning

One theory of musical development is Swanwick and Tillman's spiral model. Swanwick and Tillman (1986, as cited in Hargreaves & Zimmerman, 1992) propose a model consisting of four levels, each involving a transition in musical behavior which together form a "spiral." The first level, *mastery*, is characterized by a transition from sensory musical behavior (an exploratory approach to music, in which the learner is mainly reacting to sound and its production) to manipulative musical behavior (in which the learner gradually wields more control over their technique). In the second level, *imitation*, the behavior shifts from personal expressiveness (tending to be spontaneous and uncoordinated) to the vernacular (i.e., musical conventions). The third level is known as *imaginative play* and involves a development from speculative composition (in which students experiment with and attempt to deviate from the well-known musical conventions), to idiomatic composition (in which students are able to form those deviations into a consistent musical style). In the fourth level, termed *metacognition*, music learners transition from symbolic expression to systematic expression. Symbolic expression is more individualized, and sense of self-awareness is strong. In systematic expression, awareness is focused on the stylistic principles of the specific musical genre (Hargreaves & Zimmerman, 1992).

Another view of music learning focuses on one specific shift in the way music is categorized in the mind (Taetle & Cutietta, 2002). The early stages (or absence) of music training are characterized by a holistic or abstract categorization of musical sounds. That is, the individual elements – such as pitch, rhythm, timbre, harmony, etc. – are not factors in determining, for example, the genre to which the piece should belong (even though individual styles often have unique musical "profiles"). Instead, it is the overall perception of the piece as a whole that determines how the listener will label it (Taetle & Cutietta, 2002). Indeed, a study completed at Kent State University involving non-music majors revealed that subjects without musical training experience tended to categorize music in a holistic, rather than elemental, fashion (Booth & Cutietta, 1991). It is only as musical skills become more refined that music representations in the mind of the learner become more reflective of their individual elements (Taetle & Cutietta, 2002). Presumably, it is the exposure to highly complex auditory stimuli in the course of musical training that triggers the change in categorization strategy, though little (if any) research has been done to determine the course or specifics of this transition.

Finally, one of the more prominent music learning theories is that put forth by Edwin Gordon. Central to this model is the concept of audiation – that is, the sensation of hearing or feeling music when it is not present acoustically (Gordon, 1986, as cited in Kratus, 1994). Gordon argues that, "Sound itself is not music. Sound becomes music only through audiation, when, as with language, you translate sounds in your mind and give them meaning" (Gordon, 2003, p. 5). Gordon identifies two general types of learning skills: *discrimination* and *inference*. *Discrimination* includes imitating segments of music (e.g., played or sung by a teacher) and memorization of pieces learned from written music. Students are conscious of this type of learning, which serves as the means by which they acquire an inventory (or "vocabulary") of familiar tonal and rhythm patterns, and learn to discriminate among them (Woodford, 1996). By contrast, *inference* learning is characterized by the transfer of knowledge about familiar patterns to unfamiliar ones – that is, students infer from what is already known when presented with new information (Woodford, 1996; Gordon, 2003). Gordon acknowledges, however, that the precise way in which music students learn to make these inferences is not known (Gordon, 2003).

Word Learning and Music Learning: Similarities and Differences

As described in previous sections, there has been much interest in examining the processes of word learning and music learning. For both domains, learning can occur as a result of conscious effort (ostensive naming, imitation), or it can occur more instinctively (learning from context, inference). In both cases, learning is centered upon discovering meaning. However, "meaning" does not imply quite the same thing for language as it does for music. Understanding the meaning of a word entails mapping the mental representation of that word onto the correct linguistic form (Bloom, 2000). In music, understanding of meaning refers to the capacity to perceive and manipulate musical imagery (i.e., the ability to sense music in the absence of physical sound) in the mind (Woodford, 1996; Clark, Williamon, & Aksentijevic, 2012).

In determining meaning of either words or music, prior experience plays a role. Specifically, it has been suggested that music students draw from what they know from previously encountered material to create internal musical representations; that is, they utilize prior knowledge in examining what they are currently hearing (Bamberger, 2000, as cited in Taetle & Cutietta, 2002). Existing knowledge is also a critical element in the selective comparison component of the Sternberg and Powell (1983) theory of novel word learning. Thus it would seem that, at least theoretically, the ability to exploit prior knowledge when processing new information is important for learning both music and words.

As discussed above, it has been argued that inferring from context is the means by which the majority of words are learned – that is, the learner uses cues (value, class membership, etc.) that are drawn from the phrase, sentence, or passage in which the unknown word appears. This is not the case for music learning. For example, if a child encounters an unfamiliar piece of music, they do not learn it by inferring from the familiar pieces that they heard before and after the new one. Yet music learning is also aided by context. In the process of word learning, multiple occurrences of a novel word can aid in determining its meaning, especially if the word appears in a different context each time (Sternberg & Powell, 1983). Likewise, multiple "hearings" of the same piece of music facilitates learning. However, it is not the changing context of the music that engenders music learning; rather, it is the listener's perception of the musical composition that can change across repeated hearings (Bamberger, 1991, as cited in Taetle & Cutietta, 2002). For example, upon hearing a piece of music for a second time, a listener may notice particular subtleties that were not evident to them upon the first hearing, thus their perception of the piece is altered slightly in the process of learning.

Experience and Learning

Near transfer. The present study tested the hypothesis that musical training enhances the learning of new information. Central to this hypothesis is the assumption that musical training benefits not only music-related skills, but also performance in other non-musical disciplines. This assumption is based, at least in part, in near-transfer theory, which proposes that a concept or skill learned in one context can be applied in a similar context. It has been asserted that this transfer is an integral component of learning (Schellenberg, 2001).

According to Schellenberg (2001), the transfer of skills across contexts, domains, or modalities can be positive or negative. *Positive transfer* occurs when previous experience facilitates learning in a new context. Knowledge gleaned from past problem-solving makes it easier to solve new problems. Alternatively, previous experience could interfere with learning and produce *negative transfer*. This form of transfer can occur *proactively*, in which prior experience interferes with subsequent problem-solving, or *retroactively*, in which more recent experience makes it difficult to access the mental representations that were encoded previously (Schellenberg, 2001).

Besson and colleagues have suggested that a boost in perceptual acuity in one domain following long-term sensory experience in another domain serves as evidence of a transfer of training effects (Besson, Chobert, & Marie, 2011). This hypothesis is based on findings that certain brain structures function similarly when processing language and music (Besson et al., 2011). Results of an fMRI study on young adult non-musicians showed increased activation in a neuronal network that included portions of Broca's (Brodmann area 44) and Wernicke's (Brodmann area 22) areas for musical chord sequences containing dissonant tone clusters, compared to those containing in-key chords (Koelsch et al., 2002). Because the involvement of these areas in language processing is well established, this study appears to demonstrate an overlap in the representation of linguistic and musical syntax in the brain (Patel, 2008).

On the other hand, studies involving behavioral measures of musical and language syntax processing have demonstrated an apparent dissociation. For example, following a pair of strokes (resulting in lesions to the temporal and inferior parietal regions of the left hemisphere), composer Vissarion Shebalin exhibited deficits in speech perception and understanding, but his ability to analyze and compose music remained unaffected (Luria, Tsvetkova, & Futer, 1965). More recent studies evidence the opposite – that is, deficits in recognizing melodies, but not voices or spoken lyrics (Peretz, 1993; Griffiths et al., 1997; Ayotte, Peretz, & Hyde, 2002). In an attempt to reconcile these findings, Patel (2003) offered the hypothesis that musical and linguistic syntax do not overlap in terms of how they are *represented* in the brain (i.e., storage of predicted syntactic categories as the stimulus is perceived over time), but rather how they are *integrated* (i.e., how each new element [word or note] connects with elements that have occurred previously). Thus, the apparent shared cortical regions of activation for music and language reflect similar syntactic *integration* across domains, while behavioral results (suggesting a dissociation between musical and linguistic syntax) reflect domain-specific syntactic *representation* dysfunction (Patel, 2003).

Intersensory redundancy hypothesis. In near transfer theory, experience in one context guides learning in another. Experience also plays a central role in the intersensory redundancy hypothesis (IRH), initially put forth by Bahrick and Lickliter in 2000. Intersensory redundancy occurs when the same information is presented via multiple senses in the same place at the same time. The overlap of information provided by two or more senses (i.e., multimodal stimulation) drives perceptual, cognitive, social, and emotional development (Bahrick, Lickliter, & Flom, 2004).

A key part of the IRH is the notion that sensory information can be either: (1) *amodal*, in which multiple senses provide redundant information about the stimulus, or (2) *modality-specific*, in which only one sense can convey a particular stimulus property. Further, the sensory stimulation can be either unimodal (only one sense experiencing the stimulus property; e.g. visual) or multimodal (multiple senses experiencing the stimulus property; e.g., audiovisual). According to the IRH, multimodal perception makes amodal stimulus properties stand out compared with modality-specific properties, while unimodal perception makes modality-specific stimulus properties stand out compared with amodal properties. Thus, the IRH predicts that learning is facilitated when amodal stimuli are experienced multimodally and when modality-specific stimuli are experienced unimodally (e.g., Bahrick & Lickliter, 2000; Bahrick et al., 2004). As sensory experience accrues, perceptual context (unimodal or multimodal) becomes less of a constraint on the detection of amodal or modality-specific properties. That is, perceptual processing becomes more flexible (Bahrick et al., 2004).

In a study aimed at determining whether infants could detect a change in the orientation (a modality-specific property) of a tapping hammer, results revealed that 3- and 5-month-old infants were able to detect the change when the stimuli were presented unimodally (visually) but not when they were presented multimodally (audiovisually), consistent with the IRH prediction (Bahrick et al., 2004). However, in the same study, 8-month-old infants were also administered the orientation task and were able to detect the change in the modality-specific property in both conditions (unimodal and multimodal).

This suggests that added experience makes perceptual capability more flexible, enabling infants to pay attention to modality-specific stimulus properties whether unimodal or multimodal stimulation is available (Bahrick, Lickliter, & Flom, 2003). It is not clear how far the development of this perceptual flexibility may extend beyond infancy, though Bahrick and colleagues (2004) suggest that adults may apply this skill to novel or challenging stimuli. If this is the case, one could presume that individuals with musical training, who have had a good deal of perceptual experience in several modalities (auditory, visual, tactile) would have more flexibility for processing either type of stimulus property in either type of modality compared to individuals who lack such experience. At least one study has found that individuals with musical experience are better able to concentrate on a specific sound source while ignoring competing sources, indicating superior attention for isolating a single stimulus from multiple sources of information within the same modality (Strait, Kraus, Parbery-Clark, & Ashley, 2010). It is possible that musicians may also have an advantage for processing modality-specific information from multimodal stimuli. Such a finding would be in line with the prediction of the IRH that additional sensory experience can make perception more efficient (Bahrick et al., 2004).

Cognitive reserve. Another theory that focuses on the flexibility of neural processing is cognitive reserve. Cognitive reserve is generally considered to be the capacity of the brain to optimize performance when cognitive demand is high. Cognitive reserve may function to increase efficiency or flexibility of brain networks as a means of coping with neural pathology or particularly difficult tasks, and may function to recruit brain areas or networks not typically involved in a particular task in order to compensate for some type of insult (e.g. auditory deprivation due to hearing loss) that increases cognitive load (Stern, 2002; Stern, 2009).

It has been suggested that cognitive reserve may account for improved cognitive processing in individuals with musical experience. Hanna-Pladdy and MacKay (2011) found that auditory working memory scores increased with extent of musical experience/training in older adults. The authors point to evidence that lifestyle factors – including music activities and multilingualism – boost cognitive flexibility (a reflection of cognitive reserve), yielding more effective use of neural resources (Hanna-Pladdy & MacKay, 2011).

Enhanced cognitive reserve may also lead to superior perceptual abilities. Zendel and Alain (2012) identify cognitive reserve as a possible explanation for the correlation they observed between consistent musical practice and a preserved ability to perceive speech in noise with increasing age. They suggest that when cognitive reserve is enhanced (e.g. through musical training), sufficient cognitive resources are then available for processing auditory stimuli in adverse listening conditions (Zendel & Alain, 2012).

Stern identifies two components of cognitive reserve: (1) *neural reserve*, which is the brain's ability to cope with neural pathology or particularly difficult tasks by utilizing networks or strategies more efficiently or in a more flexible way, and (2) *neural compensation*, which is the recruitment of brain areas or networks not typically involved in a particular task, often as a means of compensating for some type of insult (traumatic injury, stroke, sensory deprivation, etc.) that increases cognitive load (Stern, 2002; Stern, 2009). While neural reserve involves more efficient use of brain networks associated with particular tasks, neural compensation consists of the recruitment of alternate networks in an attempt to make up for changes (occurring as a result of some type of neural insult) that impact the primary network (Stern, 2002; Stern, 2009). However, it has been noted that use of alternate networks tends to yield poorer task performance (Stern, 2009). It is therefore likely that the preservative effect of music training on perceiving speech in noise found by Zendel and Alain (2012) was observed not because the musicians were better at compensation, but rather because they were able to continue to use the primary network. Non-musicians, by contrast, may have had to recruit alternate neural networks, resulting in comparatively poor speech-in-noise performance. Thus, while the relation between neural compensation and music-related benefits to learning is rather ambiguous, findings such as these illustrate the support that the theory of cognitive reserve – and in particular, neural reserve – lends to the hypothesis that musical training enhances auditory learning.

Physiologic Advantages of Music Training

Pitch processing. One of the variables moderating word learning from context (identified by Sternberg and Powell, 1983) is the importance of the unknown word to the phrase in which it is embedded, as this will determine how critical it is to decipher the meaning of the word. A potential way to assess a novel word's significance in a sentence is through prosodic cues. Prosodic stress is often used to draw attention to keywords in a sentence. Prosodic information such as stress and intonation are carried by pitch contour (i.e., pattern of changes in pitch) in speech, whereas pitch contour in music conveys melody (Chandrasekaran & Kraus, 2010).

Musicians have demonstrated an advantage for detecting pitch changes in both music and speech. For example, Schön, Magne, and Besson (2004) used a task requiring recognition of incongruous pitches in music and speech to assess pitch processing in young adults with and without music training. Subjects were presented with spoken sentences and melodies and asked to determine if the final word or note was “correct” or not. One third of the stimuli (for both speech and music) were considered prosodically or melodically congruous, an additional third were weakly incongruous, and the remainder were strongly incongruous. The fundamental frequency (F0) of the final words in the weakly incongruous condition was increased by 35%; F0 was increased by 120% in the strongly incongruous condition. For music stimuli, weak incongruities were created by increasing the F0 of the final note by one fifth of a tone, while strong incongruities were created by increasing F0 by a halftone. Congruous words and notes were not manipulated. The largest difference across groups occurred for the weak incongruous condition (with musicians significantly outperforming non-musicians), indicating that musicians were better at detecting more subtle pitch violations than their non-musician peers. These results suggest that musical training provides an advantage for the perception of pitch contour in both music and spoken language (Schön et al., 2004). Thus, musical experience may enhance sensitivity to prosodic cues, allowing the listener to determine the importance of a novel word relative to the overall meaning of the sentence in which it appears.

Timing. The human auditory system is capable of encoding several temporal dimensions of sound. According to Kraus and colleagues, *timing* refers to the precise representation of these temporal features in the neural responses (Kraus, Skoe, Parbery-Clark, & Ashley, 2009). One such dimension is the temporal envelope, which represents the slow variations in the amplitude of the acoustic stimulus, and is present in both speech and music signals. When speech is masked by background noise, the noise functions to obscure the envelope of the speech. However, if the noise is also modulated, accurate encoding of the envelope of both the masker and the signal of interest can help to resolve the masker, thereby improving speech recognition (Grose, Mamo, & Hall, 2009; Peters & Hall, 1994; Takahashi & Bacon, 1992). Interruptions to temporal envelope processing may therefore explain, at least in part, why speech understanding in noise may be worse than would be predicted by audiometric thresholds. Age effects on processing of timing cues have been observed for low (Takahashi & Bacon, 1992) as well as high envelope modulation rates (Grose et al., 2009; Purcell, John, Schneider, & Picton, 2004). However, earlier studies suggest that even mild hearing loss may do more to degrade envelope processing than advancing age (Peters & Hall, 1994; Takahashi & Bacon, 1992).

Evidence of enhanced neural representation of timing suggests a potential advantage or even protective effect of musical training on temporal processing. Musacchia, Sams, Skoe, and Kraus (2007) measured auditory brainstem responses to speech and music stimuli in young adult musicians and non-musician controls. Musicians demonstrated more precise representation of timing (evidenced by earlier and larger peaks in the response to the stimulus onset) for both types of stimuli, relative to the control group (Musacchia et al., 2007). Similarly, in a study by Strait, Kraus, Skoe, and Ashley (2009), musicians showed enhanced timing in subcortical responses to vocal stimuli, compared with non-musicians. Musicians’ superior ability to encode envelope cues precisely may mean that they are more proficient at processing the acoustic features of both familiar and novel words, especially when they are encountered in difficult listening conditions.

Timbre. Accurate detection of timbre helps listeners discriminate between two distinct sounds with identical pitch (Kraus et al., 2009). In speech, timbre carries acoustic information that aids in distinguishing one phoneme from another. Timbre is considered a qualitative acoustic characteristic, and is dependent upon harmonic components of complex sounds (Lee, Skoe, Kraus, & Ashley, 2009).

Musicians have been found to differ from non-musicians in their sensitivity to timbre. Lee et al. (2009) compared brainstem responses to musical intervals (i.e., two tones of differing pitch presented sequentially) in two groups of young adults: those with 10 or more years of music training (“musicians”), and those with fewer than 3 years of music training (“non-musicians”). Relative to non-musicians, the musician group showed stronger physiologic responses to the harmonic components of the higher of the two tones. This suggests that the auditory systems of individuals with musical training experience are highly tuned to timbre, at least at the level of the brainstem (Lee et al., 2009). Using similar subject groups, Musacchia, Strait, and Kraus (2008) examined brainstem responses to the speech syllable /da/ (F0 = 100 Hz). Response amplitudes for the harmonic components of the speech stimulus (i.e., 200 Hz, 300 Hz, etc.) were higher in musicians than in non-musicians. These results indicate that neural representation of timbre-related aspects of speech is enhanced in musicians (Musacchia et al., 2008). As the perception of timbre is important for the discrimination of speech sounds, heightened sensitivity to timbre could improve detection of novel words that differ from familiar words by only one phoneme.

Speech in noise. Compared to non-musician peers, musicians have demonstrated superior speech perception in noise. For example, in a 2009 study, Parbery-Clark, Skoe, Lam and Kraus tested young adult musicians and non-musicians on two clinical speech-in-noise (SIN) assessments. For both measures, the musician group scored significantly higher than the non-musician group (Parbery-Clark et al., 2009b). Additionally, speech-evoked auditory brainstem responses (ABRs) of musicians have shown less degradation with the addition of background noise compared with those of non-musicians. In a study of young adult musicians and non-musicians, evoked responses to SIN were measured and compared to responses to speech in quiet. Musicians' responses in the two conditions showed a higher correlation (in terms of morphology) compared to those of non-musicians (Parbery-Clark, Skoe, & Kraus, 2009a). Similar results have been obtained in children. Strait, Parbery-Clark, Hittner, and Kraus (2012) studied two groups of school-age children (mean age = 10.2 years): 15 children enrolled in music training since age 5 ("musicians"), and 16 children with fewer than 5 years of musical training ("non-musicians"). SIN perception was measured both behaviorally, using standardized measures of speech recognition in noise, and electrophysiologically, using speech-evoked ABR. Again, subjects with more musical experience scored higher on behavioral measures and showed less degradation of evoked responses when noise was added to the speech stimulus. In contrast to the results found in adults, however, child musicians also demonstrated earlier peak latencies on speech-evoked ABR in both quiet and noise, indicating superior encoding of speech harmonics (Strait et al., 2012).

One possible explanation for the musical advantage in SIN perception has to do with auditory scene analysis (that is, the process of segregating the signal of interest from competing signals), on which perception of SIN is thought to rely (Parbery-Clark et al., 2009b). Sensitivity of the auditory system to changes in the auditory scene can be measured using mismatch negativity (MMN), an ERP occurring 100-150 milliseconds after stimulus onset. MMN is evoked using an oddball paradigm, involving the presentation of a standard stimulus (high probability) and a deviant stimulus (low probability) that differ from each other on some parameter (e.g., frequency, level, duration). The evoked response to the deviant shows a larger negative peak than the response to the standard. This response is considered evidence of the sensitivity of the auditory system to changes in the auditory scene (Schnupp, Nelken, & King, 2011). Musicians have shown larger MMN responses to changes in melodic contour and interval compared with their non-musician peers, suggesting enhanced automatic encoding of these specific acoustic features (Pantev et al., 2003). This implies that – at least physiologically – musicians are better able to detect changes in the auditory scene than non-musicians, possibly making them better equipped to perceive speech in adverse listening conditions (Parbery-Clark et al., 2009a).

Superior processing of temporal fine structure (TFS) provides a second potential explanation for observations of better SIN perception in musicians compared to non-musicians. In contrast to the temporal envelope (i.e., slow modulations in envelope amplitude), TFS represents the rapid changes in the amplitude of the acoustic stimulus. In speech, TFS cues allow consonants to be distinguished from one another, and provide information regarding vowel formant transitions (Rosen, 1992). Previous studies (e.g., Lorenzi, Gilbert, Carn, Garnier, & Moore, 2006) indicate that TFS may be particularly important for speech recognition in fluctuating noise. In a phenomenon known as ‘‘listening in the dips,” normal-hearing listeners are able to “glimpse” the fine structure of the signal in the temporal valleys of the fluctuating noise; this fine structure information aids in speech perception (Moore, 2008; Brown and Bacon, 2010). Sensitivity to TFS has been examined in musicians and non-musicians using a two-interval forced-choice task in which listeners discriminated between harmonic and inharmonic complex tones (Mishra, Panda, & Raj, 2015). Inharmonic tones were created by shifting the harmonics of the harmonic tone upward in frequency. Musicians were able to detect smaller frequency shifts than non-musicians, suggesting that musicians show enhanced sensitivity to TFS (Mishra et al., 2015). It is possible that this enhanced fine structure sensitivity leads to greater “dip listening” acuity in musicians compared to non-musicians, thereby giving musicians an advantage on SIN perception.

Working memory. Working memory (WM) refers to the short-term storage and manipulation of information (Baddeley, 1992). WM plays a key role in higher-level cognitive tasks, including learning and speech processing (Schulze & Koelsch, 2012; Baddeley, 1992). Previous studies have found that musically-trained adults and children exhibit greater WM capacity. Adult musicians and non-musicians tested in the speech-in-noise study detailed in the previous section (Parbery-Clark et al., 2009b) were also administered the Numbers Reversed and Auditory Working Memory subtests of the Woodcock-Johnson III Cognitive test (Woodcock, McGrew, & Mather, 2001). The Numbers Reversed test requires subjects to repeat a list of numbers in backward order, while the Auditory Working Memory test involves mixed sets of words and numbers, which the subjects must separate and list in sequential order. Scores on these tests were combined to yield a composite measure of WM. Results revealed significantly greater

WM capacity in musicians compared to non-musicians (Parbery-Clark et al., 2009b). The Auditory Working Memory subtest was also used to assess auditory WM in the child musicians and non-musicians from the speech-in-noise study discussed above (Strait et al., 2012). Again, the musician group performed significantly better on this measure than their non-musician peers (Strait et al., 2012). In a longitudinal study, the backward digit recall subtest of the Automated Working Memory Assessment (Alloway, 2007) was used to evaluate verbal WM in children and young adults between the ages of 6 and 25 years (Nutley, Darki, & Klingber, 2014). Subjects repeated sequences of numbers in reverse order, starting with only 2 numbers and increasing until the subject failed to repeat all of the digits in the correct order. Musical experience was a significant predictor of performance on this task, suggesting that activities like music, which place demands on WM, may strengthen WM capacity (Nutley et al., 2014).

Pragmatics. There is evidence that working memory capacity may be related to variations in pragmatic abilities. According to Baltes (1993), cognitive pragmatics include skills such as reading, writing, and language comprehension, as well as more abstract abilities such as self-awareness and life skills. In a study by Hannon and Daneman (2009), younger (18-25 years old) and older (64-87 years old) adults were administered a test of reading comprehension (Nelson-Denny, Form E; Brown, Bennett, & Hanna, 1981), in which participants read 8 short passages, then answered 36 multiple-choice questions regarding the content of these passages. The older adults showed poorer reading comprehension than their younger counterparts. However, working memory ability (as measured by a reading span test in these same participants) accounted for 44%

of the variance in reading comprehension scores in older adults, whereas it accounted for only 22% of reading comprehension variance among younger adults.

In an earlier study involving only young adults, Daneman and Carpenter (1983) used an alternate reading comprehension task, in which participants read passages containing ambiguous words whose meanings could be inferred from context. Participants also completed the same reading span test used in the Hannon and Daneman (2009) study, which required them to read aloud sets of sentences (2-6 sentences each), then recall the last word of each sentence in the set. Participants with higher working memory scores performed better on the reading comprehension task, indicating that these individuals were better at inferring meaning from context (Daneman & Carpenter, 1983), a task that requires the ability to draw from existing knowledge and life experience.

Music experience may also have the potential to enhance pragmatics. Corrigall and Trainor (2011) tested reading comprehension in 46 children between 6 and 9 years of age, all of whom participated in music lessons. Musical experience was quantified as the number of years of training as well as the age of training onset. Participants completed the Passage Comprehension subtest of the Woodcock Reading Mastery Test – Revised (WRMT-R; Woodcock, 1987), which required them to identify a missing word in a sentence or short passage. Essentially, this test assesses ability to use context and experience to generate the missing information that would make each statement comprehensible. Reading comprehension scores were significantly correlated with the number of years of music training and with the age at which music training began. These correlations remained robust even after controlling for multiple factors including age, general intelligence, and number of hours spent reading per week (Corrigall & Trainor, 2011). Thus, auditory experience in the form of music training may facilitate reading comprehension (a pragmatic skill), either directly or via enhancements to other cognitive functions.

CHAPTER 3

CONCEPTUAL FRAMEWORK

The relationship between auditory experience and auditory performance (behavioral and electrophysiologic) was investigated in the context of a model put forth by Baltes (1993). Baltes posits a dual-process categorization of intelligence that contrasts cognitive mechanics with cognitive pragmatics. *Cognitive mechanics* consist of the neurophysiologic structure and processing of the brain. They reflect mechanisms that emerged as a result of evolution, including the speed and accuracy of detection and categorization of sensory information. By contrast, knowledge and skills based in culture, which result from participation in society, comprise *cognitive pragmatics*. These reflect knowledge about the world and the self. According to the Baltes model, cognitive mechanics may be quantified with objective physiologic and behavioral measures (e.g., Strait et al., 2012; Bidelman et al., 2014a; Bidelman & Alain, 2015), whereas cognitive pragmatics are reflected by language knowledge and comprehension (Hannon & Daneman, 2009; Daneman & Carpenter, 1983).

Novel word learning involves aspects of both components of this intelligence model. Cognitive mechanics are necessary for acoustic-phonetic pattern recognition, a process triggered by auditory input. Higher-level lexical knowledge (cognitive pragmatics) allows for the detection of the unknown word as novel information, which initiates learning. In a process known as *configuration,* acoustic/phonetic (bottom-up) and lexical (top-down) inputs are combined to form a cognitive representation of the novel word within the lexicon (Pittman & Rash, 2016). In other words, the ability to map sensory information onto a single specific construct, excluding others, requires: (1) the ability to segregate and organize acoustic features of the speech input, and (2) the ability to use context and prior knowledge to apply an appropriate set of rules to the novel information to form a lexical representation. This suggests that cognitive mechanics and cognitive pragmatics each play a role in novel word learning.

According to the Baltes model shown in Figure 1, cognitive mechanics develop early in life, then weaken with advancing age. For example, older adults show degraded auditory brainstem responses to speech stimuli compared to younger adults (Bidelman et al., 2014a). Cognitive pragmatics build upon mechanics and are an accumulation of experience and knowledge gained across the life span (Cole, 1996). Baltes argues that, in contrast to cognitive mechanics, cognitive pragmatics are not susceptible to age-related decline. There is evidence, however, that pragmatic performance may indeed decrease with age (Hannon & Daneman, 2009), though the onset of deterioration may occur later in life and its trajectory may be more gradual than that of cognitive mechanics.

Added to the figure is a function representing learning, which, like cognitive mechanics, begins to develop early in life, but like cognitive pragmatics, is relatively well preserved in late adulthood (Pittman et al., 2017). This addition illustrates the contribution of cognitive mechanics, which is the bottom-up processing of, for example, the acoustic pattern of a new word, while pragmatics provides the ability to map this sensory information onto a single specific construct. Pragmatics also includes the ability to use context and prior knowledge to apply an appropriate set of rules to the novel input to form a meaningful representation. Because learning influences and is influenced by cognitive mechanics and cognitive pragmatics, the processes involved in learning may be common to both components. The theoretical framework of this study expanded Baltes's model to include a component of cognitive refinement that represents the ongoing process of learning new information. The objective of this study was to examine the impact of auditory experience on auditory performance using assessments that reflect each of the three components of this expanded model.

CHAPTER 4

HYPOTHESES

The central hypothesis of this study was that enhancements to cognitive mechanics due to enriched auditory experience consequently improve learning capacity and pragmatic ability. To test this hypothesis, each of the three components of the expanded Baltes model (mechanics, learning, and pragmatics) was assessed in a group of young normal-hearing adults with a history of highly complex auditory experiences in the form of long-term music training. Their performance was compared to that of a group of young adults lacking any musical experience. *Mechanics* were quantified using physiologic measures of cortical activity that reflect speech sound processing. *Learning* was defined as the speed with which participants are able to map nonsense words onto novel images. *Pragmatics* was determined by measuring participants' ability to consolidate newly-learned information into long-term knowledge and recognize this information in a different context. Participants with enriched auditory experience (i.e., musical training) were expected to exhibit shorter wave latencies and larger peak-to-peak amplitudes on auditory evoked potentials, and to learn novel words at a faster rate than participants without musical training. It was also expected that participants with musical experience would be able to more effectively consolidate the new words into long-term memory, allowing them to both recall and detect newly-learned words more accurately over time. Finally, it was anticipated that group differences would be larger for both behavioral and electrophysiologic performance when stimuli were presented in multi-talker babble compared to quiet, yielding significant group x listening condition interactions.

CHAPTER 5

METHODS

Cognitive *mechanics* were assessed by measuring cortical auditory evoked responses to speech stimuli. *Learning* ability was determined using a rapid word-learning paradigm for nonsense words (Pittman, 2008; Pittman, 2011). Finally, *pragmatics* were measured by testing participants' retention of learned nonsense words, as well as their ability to detect them within continuous discourse. Additionally, all 3 tasks were completed in quiet and in a background of multi-talker babble to examine the impact of a degraded auditory signal on mechanics, learning, and pragmatics. The babble condition was included as a means of determining whether enhanced auditory experience has a protective effect against the consequences of a type of auditory distortion similar to that experienced by individuals with hearing loss.

Participants. Twenty adult musicians between the ages of 23 and 34 years were recruited from the student population in the School of Music at Arizona State University (ASU) and from local community symphonies. Musical history was characterized in terms of the age when musical training was initiated and years of consistent practice. Musicians had between 13 and 25 years of training (mean=18.8 ±4.0 yrs), initiated between 3 and 12 years of age (mean=7.7 ±2.6 yrs). All participants held at least a bachelor's degree. For 19 of the 20 musicians, this degree was in a music-related discipline (performance, education, theory, etc.), whereas one musician held a degree in a non-music discipline.

An additional 20 adults between the ages of 21 and 38 years with little to no musical experience served as a control group. Ten of the non-musicians had no music training at all, while the other 10 had between 0.25 and 2.25 years of training (mean=0.5 ±0.7 yrs), initiated between ages 6 and 15 years (mean=11.0 ±2.8 yrs). None of the non-musicians reported any training within the past 10 years.

All participants had normal hearing bilaterally, as confirmed by pure tone audiometry. Hearing thresholds were ≤25 dB HL for octave frequencies between .25 and 8 kHz, with the exception of one musician, whose threshold at 8 kHz was 30 dB HL in the right ear only. Average hearing thresholds are shown in Table 1 along with demographic information for each participant group.

Mechanics. Cognitive mechanics were assessed using cortical auditory evoked potentials (CAEPs). The components of interest were N1, P2, and P3a. The N1-P2 complex amplitude has been shown to increase following short-term (10 days) speech sound training, compared to pre-training results (Tremblay, Kraus, McGee, Ponton, & Otis, 2001). Larger N1-P2 amplitudes have also been observed in listeners with long-term musical experience compared to non-musicians, suggesting that this component changes with auditory experience (Shahin, Bosnyak, Trainor, & Roberts, 2003; Bidelman, Weiss, Moreno, & Alain, 2014b). The P3 component is thought to reflect attention and memory (e.g., Bidelman and Alain, 2015) – functions that play a critical role in learning. The P3 wave may also be bifurcated into two subcomponents: (1) *P3a*, which is produced in response to an oddball auditory stimulus without a task, and (2) *P3b*, which is a task-related potential observed when the subject is actively attending to the stimuli (Squires, Squires, & Hillyard, 1975; Snyder & Hillyard, 1976). P3a latency in particular has been found to be negatively correlated with working memory scores (Polich, Howard, & Starr, 1983).

Stimuli. The evoking stimuli consisted of five 100-msec synthetic vowel sounds on a perceptual phonetic continuum from /u/ to /a/. Vowel tokens differed in the frequency of the first formant (F1), while all other formants (fundamental [F0], F2, and F3) were identical across tokens (see Table 2 for formant frequencies). Stimuli were organized into two sets: a *wide* set, consisting of the tokens whose F1 frequencies were furthest apart (430, 585, and 730 Hz), and a *narrow* set, consisting of the tokens whose F1 frequencies were close together (490, 585, and 660 Hz). For the babble condition, a 6-talker (3 male, 3 female) multi-talker babble was used. Multi-talker babble was routed from the computer through a low noise (SNR 111 dB), high speed (96 kHz), high quality (24-bit resolution) soundcard with 8 analog channels (RME Multiface II soundcard) through a standard clinical audiometer (GSI 61; Grason-Stadler) to an insert earphone (ER-3A; Etymotic Research) and presented at 60 dB SPL.

Equipment. Stimuli were presented at 70 dB SPL to one of the participant's ears using SmartEP (Intelligent Hearing Systems). The same equipment was used to collect the evoked responses. Stimuli were routed through an insert earphone and placed in the same ear as the insert routed from the sound card.

Procedure. Six disc electrodes were placed on the participant's head in a vertical montage, with the active electrode at the vertex (Cz), the ground electrode on the high forehead (Fpz), and reference electrodes on right (M2) and left mastoids (M1). Right and left mid-coronal electrodes were also placed at C4 and C3, respectively. Responses were recorded in four channels: Cz-M2, Cz-M1, C4-M2, and C3-M1. Contact impedance for all electrodes did not exceed 7 kΩ. Participants were seated in a chair in a sound-treated room and were instructed to remain relaxed but awake with eyes open during the test. To facilitate this, participants were permitted to look at their smartphone or tablet computer during this portion of the testing.

For each stimulus set (wide, narrow), the three vowel tokens were presented in random order to one of the participant's ears. Within each set, the token with the highest F1 was the standard, which accounted for 60% of the total presentations, while the middle and lower F1 tokens were the two deviants, each of which accounted for 20% of the presentations (see Table 3). CAEPs were collected until approximately 75 artifact-free responses to the each of the deviants were obtained. Ear (right, left), set (wide, narrow), and listening condition (quiet, babble) were counterbalanced across participants.

Data reduction. The following peaks were identified for each participant from the resulting waveforms for each ear, vowel set, and listening condition: N1 (70-200 ms), P2 (140-250 ms), and P3a (240-400 ms). The latencies of each of these components were measured, as well as the magnitudes of the N1-P2 complex and P3a component.

Categorical perception. Participants also categorized vowel tokens as either /u/ or /a/ by selecting the appropriately labeled button on a computer screen (see Figure 2). Each token was presented 50 times for a total of 250 randomized presentations. Stimuli were presented though binaural insert earphones (ER-3A; Etymotic Research) at 60 dB SPL. Listeners were told that some tokens would be easy to categorize while others would be more ambiguous, and were instructed to select the speech sound category closest to each token. Visual feedback was provided via computer game following each selection, as an indication that the response had been recorded. Feedback was provided for each response regardless of accuracy. For each of the five vowel tokens, the proportion of trials (out of 50) for which the listener selected the /a/ category was calculated (0-100% of tokens labeled as /a/). These values were used to create categorical perception functions across the vowel continuum (/u/ to /a/).

Due to equipment limitations, the categorical perception task could not be completed during CAEP recording. Thus, this task was administered the day preceding evoked potential testing. As reaction time could not be recorded during categorical perception, scores on this task could not be calculated in precisely the same way as in previous studies (e.g., Bidelman et al., 2014a). Therefore, the slope of the psychometric function was calculated between the 2nd and 4th vowel tokens (F1=490Hz, 660Hz) to quantify the listeners' ability to distinguish between speech sounds based on variation in F1 frequency.

Learning. To assess multiple stages of word learning, participants completed a battery of behavioral tasks involving familiar and unfamiliar word stimuli. The rationale for using a set of tasks of varying difficulty was to identify more precisely the level of cognitive load at which musical experience becomes advantageous.

Word recognition. Listeners heard and repeated aloud sets of 25 familiar words from the NU-6 word-recognition test (Tillman & Carhart, 1966). Participants' verbal responses were captured with a digital audio recorder (Olympus, WS 801/802) coupled to a head-worn microphone (Shure, WH20) positioned ~2 inches from the corner of the speaker's mouth. Responses were scored offline as either correct or incorrect. Overall performance was scored in percent correct. No reinforcement was provided for this task.

Non-word detection. Participants heard lists of 12 sentences each containing 4 monosyllabic words and indicated which words (if any) in each sentence were nonsense words by selecting the corresponding numbered button(s) on a computer screen. Within each list, 6 sentences contained two nonsense words, 4 sentences contained one nonsense word, and 2 sentences contained zero nonsense words. Visual reinforcement was provided via an interactive computer game for correct responses, but not for incorrect responses. A screenshot of the computer game used for the non-word detection task is shown in Figure 3.

Performance on this task was assessed in a number of ways. First, the overall number of sentences in which all four words were correctly identified as either real or nonsense was summed and divided by the total number of sentences in the list (12), yielding a percent correct score for sentences. The total number of words correctly identified within the entire list of sentences was also calculated and divided by the total number of words (48), yielding a percent correct score for words. Finally, participants' scores were broken down into hits (correctly identified nonsense words), misses (nonsense words identified as real), false alarms (real words identified as nonsense), and correct rejections (correctly identified real words). Hits and false alarms were then used to calculate *d*-prime, which served as an indication of the listener's sensitivity to unfamiliar (novel) words.

Word learning. Finally, an interactive word-learning task was used to determine short-term word-learning ability. This particular paradigm has been used previously to identify listener characteristics that impact word-learning ability, including age, receptive vocabulary, and hearing level (Pittman, Lewis, Hoover, & Stelmachowicz, 2005; Pittman, 2011). Participants heard sets of five nonsense words, presented one at a time. Each nonsense word was paired with one of five nonsense objects/characters. A picture of each nonsense image was displayed on one of 5 buttons arranged across the bottom of a computer screen. Listeners selected one of the five images after the presentation of each nonsense word. Visual reinforcement for correct selections was provided by a computer game that appeared just above the row of buttons. The game advanced (e.g., one piece of a puzzle appeared) following each correct response, while no reinforcement was provided for incorrect selections. Participants were instructed to use the reinforcement to associate each nonsense word with the correct image. Each nonsense word was presented 18 times for a total of 90 randomized trials. A screenshot of the interactive computer game used for the word-learning task is shown in Figure 4.

Scores on this task were quantified in terms of percent correct (out of 90 trials) as well as word learning speed. To calculate speed of word learning, trial-by-trial data were reduced chronologically to 9 bins of 10 trials each, and the proportion of correct responses within each bin was calculated. The raw data were then smoothed with an exponential growth function $P_c = 1 - 0.8^{e-\frac{n}{c}}$, where P_c is the probability of a correct answer, 1 – 0.8 reflects chance performance for this task (20%), *e* is 2.718. . ., *n* is the midpoint of the trial block (5, 15, 25, etc.), and *c* is the time constant of the process. When the number of trials equals the time constant (*n*=*c*), performance is approximately 70% correct. The number of trials was log transformed and limited to no more than 1,000 trials. The inverse of the number of trials required for each participant to reach this criterion level of performance is the speed of word learning.

Auditory stimuli. Participants completed each task twice: once in quiet and once in the same multi-talker babble used for CAEP testing at +3 dB SNR. Order of listening condition was counterbalanced across participants. Word and sentence stimuli were recorded by a talker having a standard American-English dialect at a sampling rate of

22.05 kHz using a microphone with a flat frequency response to 10 kHz. Stimuli were digitized and edited into individual.wav files using Adobe Audition v1.5 and equated for RMS level.

Equipment. Stimuli were presented from a desktop computer using custom laboratory software. Auditory stimuli (NU-6 words, sentences, nonsense words) were routed from the computer through a high speed (96 kHz), high quality (24-bit resolution) soundcard with 6 analog channels (Echo Gina 3G) to binaural insert earphones (ER-3A; Etymotic Research) and presented at 60 dB SPL. The experimental software was also used to display visual reinforcement and record participants' responses.

Pragmatics. Cognitive pragmatics testing was conducted on the day immediately following the learning assessments. Participants completed two measures of pragmatics: a learned-word retention task, and a learned-word detection task. Learned-word detection was assessed in quiet and in a background of the multi-talker babble presented during the CAEPs and word-learning tests, and the order of the listening conditions was counterbalanced across participants.

Learned-word retention. This test was completed as a means of quantifying participants' capacity for consolidating newly-learned information into long-term memory and integrating it with existing knowledge. Participants completed a post-test consisting of the same two sets of 5 images used in the word-learning tasks completed the previous day, along with a word bank containing the 10 learned nonsense words, plus 10 foils (unfamiliar but similar nonsense words). Participants were instructed to label each image with the correct word from the word bank. Responses were scored in percent correct. Appendix A contains the post-test and word bank.

Learned-word detection. This task was designed to assess participants' ability to detect learned nonsense words within continuous discourse. For each listening condition (quiet, babble), participants listened to a spoken passage read aloud by the same talker who produced the nonsense words in the word-learning task. Some words in the passage were replaced with the learned words, while others were replaced with unfamiliar nonsense words (foils). Both passages contained 248 words (including nonsense targets and foils) that resulted in recordings that were approximately 90 seconds in duration. The passages were presented in the same listening conditions (quiet, babble) in which the words were presented in the word-learning task. Listening condition order was counterbalanced across participants. Appendix B contains a transcript of the passages containing the learned nonsense words (boldface) and their foils (italicized).

Participants were given a clicker and were instructed to click as soon as they heard a nonsense word that they learned in the previous day's task, and to ignore unfamiliar nonsense words. Three of the five learned words appeared in each passage; three repetitions of each learned word occurred within the discourse. Each passage also contained three repetitions of each of three unfamiliar foils. Participants' responses were recorded by the examiner for later analysis. The number of hits (responses to learned nonsense words) and false alarms (responses to unfamiliar nonsense words) were used to calculate *d*-prime for each participant. This value served as an indication of the listener's sensitivity to the learned words.

Participants' audible clicks were recorded using the surround microphone setting on a digital recorder (Zoom H2N), while the passage (i.e., the output of the soundcard) was recorded simultaneously in a separate channel using a hardwired input connecting directly to the soundcard. These recordings were used to verify examiner scoring.

Analysis. To test the main hypothesis, repeated measures analyses of variance were used to examine group differences in mechanics, learning, and pragmatics across listening conditions. The between-subjects factor was musical experience (long-term training, no training), while the within-subjects factor was listening condition (quiet, babble). Correlational analyses were conducted to determine the relationship between performance on the behavioral tasks and the CAEPs.

CHAPTER 6

RESULTS

Mechanics

Because the waveforms recorded in the right and left mastoid channels (Cz-M2, Cz-M1) contained peaks that were better defined than those from the right and left mid-coronal channels (C4-M2, C3-M1), responses from the mastoid channels were analyzed. Results for the three latency (N1, P2, P3a) and two amplitude (N1-P2, P3a) components were collapsed across ear, channel, stimulus set (wide, narrow), and stimulus type (standard, deviant 1, deviant 2). Main effects of group (musician, non-musician) and listening condition (quiet, babble) were examined as well as interactions for each CAEP component.

The upper panels of Figure 5 show the individual and average latencies (left) and amplitudes (right) for musicians and non-musicians collapsed across listening condition. Overall, latencies were longer for the musicians than for the non-musicians although the largest difference was found for P2 (7 msec) with smaller differences between groups for the N1 and P3a components (1-2 msec). Amplitudes were larger in the musicians than the non-musicians for the N1-P2 (0.8 µV) and P3a (0.1 µV) components. The lower panels of Figure 5 show the individual and average latencies (left) and amplitudes (right) for the quiet and babble conditions collapsed across group. Component latencies were 4-7 msec longer in babble compared to quiet, while amplitudes decreased in babble by 3-6 µV relative to quiet.

Results of repeated measures analyses of variance are shown in Table 4. Overall, music training did not have a significant effect on CAEP components. Response latencies and amplitudes were equivalent across group with the exception of P2 latency, which was unexpectedly longer in musicians than in non-musicians. The musicians did show notably larger N1-P2 component amplitudes than the non-musicians, though this difference did not reach significance. Thus, musicians did not demonstrate superior mechanics compared to non-musicians, as stronger evoked responses are generally characterized by shorter rather than longer latencies and larger component amplitudes. In contrast to group results, the effect of listening condition was significant for all components. N1, P2, and P3a latencies were significantly longer and N1-P2, P3a amplitudes were significantly smaller in babble than in quiet. Finally, no significant group x listening condition interaction was observed. That is, music training did not yield stronger evoked responses (i.e., shorter latencies, larger amplitudes) in either favorable or unfavorable listening conditions.

In previous studies, musicians' electrophysiologic responses have shown less degradation in noise than those of non-musicians (Parbery-Clark et al., 2009a; Strait et al., 2012). To determine whether similar results occurred in the present study – that is, whether musicians' responses were less impacted than those of non-musicians by the poorer listening conditions caused by multi-talker babble – separate repeated measures analyses of variance were conducted in musicians and in non-musicians. For each group, evoked response latencies and amplitudes (dependent variables) were compared across listening condition (independent variable). The top panels of Figure 6 show individual and average evoked response latencies (left) and amplitudes (right) as a function of

CAEP component for musicians in quiet and babble. For the musicians, latencies increased 4-7 msec in babble compared to quiet, while N1-P2 amplitude decreased 0.4 µV in babble relative to quiet. P3a amplitudes, however, decreased by only 0.1 µV between the quiet and babble listening conditions. The lower panels of Figure 6 show CAEP component latencies and amplitudes for non-musicians. Like the musicians, the non-musician's latencies increased 4-7 msec in babble relative to latencies in quiet, N1-P2 amplitude decreased 0.8 µV in babble, and P3a amplitudes decreased by 0.3 µV.

Results of repeated measures analyses of variance for musicians and non-musicians are shown in Table 5. For both musicians and non-musicians, P2 latencies were significantly longer in babble than in quiet. Additionally, N1-P2 amplitudes were significantly larger in quiet than in babble for both groups. In non-musicians, P3a amplitude decreased significantly in babble relative to quiet. By contrast, P3a amplitudes did not differ across listening condition in musicians. These results are consistent with previous findings showing less degradation in noise (babble) relative to quiet in musicians compared to non-musicians, at least for this later cortical auditory evoked potential (Parbery-Clark et al., 2009a; Strait et al., 2012).

Figure 7 shows the psychometric functions associated with the /u/ to /a/ vowel continuum for the musicians (left panel) and the non-musicians (right panel). The parameter in each panel is listening condition. The perceptual boundaries did not appear to shift sufficiently to indicate that listening condition or musical training had an impact on the ability to perceive small variations in vowel pitch. Figure 8 shows the individual and average slopes of the categorical perception task as a function of listening condition for each group. Repeated measures analysis of variance revealed no difference in slope for either group ($F(1,38)=0.19$; $p=0.665$, $\eta^2=0.005$) or listening condition ($F(1,38)=0.08$; $p=0.776$, $\eta^2=0.002$). The group x listening condition interaction was also non-significant ($F(1,38)=0.08$; $p=0.776$, $\eta^2=0.002$).

Learning

Figure 9 shows the individual and average word recognition as a function of listening condition for musicians and non-musicians. For both groups, performance was 11-12% poorer in babble than in quiet. Musicians' scores were also 1-2% higher than those of non-musicians in both listening conditions. Repeated measures analyses of variance revealed significantly better recognition in quiet than in babble ($F(1,38)=100.17$; $p<0.001$, $\eta^2=0.725$). However, no effect of group was found ($F(1,38)=3.52$; $p=0.068$, $\eta^2=0.085$) indicating that word recognition was equivalent across groups when collapsed across listening condition. Further, the group x listening condition interaction was not significant ($F(1,38)=0.06$; $p=0.814$, $\eta^2=0.001$), indicating that performance in both groups was similarly impacted by babble.

Figure 10 shows the individual and average performance on the non-word detection task as a function of listening condition for musicians and non-musicians. Performance is expressed in units of *d*-prime, which denotes listener sensitivity to novel (nonsense) words. For both groups, *d*-prime scores were significantly higher in quiet than in babble ($F(1,38)=134.09$; $p<0.001$, $\eta^2=0.779$), indicating better sensitivity to nonsense words in quiet compared to the babble condition. However, no main effect of group was observed ($F(1,38)=0.06$; $p=0.808$, $\eta^2=0.002$). Finally, musicians performed better than non-musicians in quiet, while non-musicians performed better than musicians in babble. However, no group x listening condition interaction was present ($F(1,38)=1.20$; $p=0.281$, $\eta^2=0.031$).

Figure 11 shows individual and average speed of word learning as a function of condition for musicians and non-musicians. For this task, performance did not decline in multi-talker babble relative to quiet ($F(1,38)=0.42$; $p=0.522$, $\eta^2=0.011$). Word learning speed in non-musicians was faster than that of musicians in quiet, while musicians learned nonsense words faster than non-musicians in babble. Even so, no main effect of group ($F(1,38)=0.34$; $p=0.565$, $\eta^2=0.009$) or group x listening condition interaction ($F(1,38)=2.84$; $p=0.100$, $\eta^2=0.069$) was observed.

Taken together, behavioral performance in multi-talker babble was expected to be poorer than performance in quiet. This was supported by the data, but only for word recognition and non-word detection, which required perception of both real and nonsense words. The rapid learning of nonsense words was not affected by the presence of multi-talker babble. Further, it was hypothesized that musicians would be able to detect unfamiliar words more accurately and learn them more quickly, owing to their enriched auditory experience. However, the data did not support this, as performance on word recognition, non-word detection, and word learning did not differ significantly across groups. Finally, it was anticipated that musicians' performance would be less impacted by noise than that of their non-musician peers, yielding a significant group x listening condition interaction. This hypothesis was not supported by the data.

Pragmatics

Figure 12 shows individual and average performance on the learned-word retention task as a function of listening condition for musicians and non-musicians. Both groups performed similarly across conditions, recalling 61-69% of nonsense words learned the previous day, on average. Performance differed across groups by 3 to 8% with non-musicians performing better in quiet and musicians performing better in babble. Repeated measures analysis of variance revealed no significant effect of group ($F(1,38)=0.07$; $p=0.789$, $\eta^2=0.002$) or listening condition ($F(1,38)=0.62$; $p=0.437$, $\eta^2=0.016$) for this task. The group x listening condition interaction also was not significant ($F(1,38)=0.58$; $p=0.449$, $\eta^2=0.015$). These results indicate that musicians did not show an advantage over non-musicians for recalling nonsense words learned the previous day.

Figure 13 shows individual and average performance on the learned-word detection task as a function of listening condition for musicians and non-musicians. Performance is expressed in units of *d*-prime to indicate the listeners' sensitivity to the newly learned words in the context of familiar words. Overall, sensitivity to the learned nonsense words was not significantly higher in quiet than in multi-talker babble ($F(1,38)=1.33$; $p=0.256$, $\eta^2=0.035$). Additionally, the performance of the musicians and non-musicians was equivalent ($F(1,38)=0.12$; $p=0.727$, $\eta^2=0.003$). Finally, no group x listening condition interaction was observed ($F(1,38)=0.01$; $p=0.925$, $\eta^2<0.001$). These results suggest that musical training experience did not improve or degrade the ability to detect newly-learned nonsense words in quiet or in babble.

Correlations between CAEPs and behavioral performance

Correlational analyses were conducted to determine the relationship between performance on the behavioral tasks and the CAEPs, and between performance on behavioral tasks completed across the two days of testing. Results of the learned-word retention and learned-word detection tasks in quiet and babble were collapsed across groups and subjected to bivariate correlation analyses with learning speed to demonstrate that performance on the learning task produced memory traces that could be retrieved the next day in an alternate modality (written post-test) and in an alternate context (spoken passages). Learned-word retention was significantly correlated with learning speed in both the quiet (r=0.58, p<0.001) and multi-talker babble (r=0.61, p<0.001) listening conditions. Learned-word detection in quiet and in babble were also significantly correlated with learning speed in quiet (r=0.41, p=0.003) and babble (r=0.36, p=0.024) conditions, respectively. These results suggest that the faster and more accurately participants learned the nonsense words, the more accurately they were able to recall and detect them in context the following day. Further, strong correlations between the speed at which listeners learned nonsense words and the accuracy with which they were able to recall and detect those words the next day suggests that the rapid word learning task used in the present study is a valid measure of learning.

The results for the five CAEP components (N1, P2, and P3a latencies; N1-P2 and P3a amplitudes) and the five behavioral tasks (word recognition, non-word detection, word learning, learned-word retention, and learned-word detection) were collapsed across groups and subjected to bivariate correlation analyses to determine if a relationship exists between the physiologic and behavioral measures. The results are shown in Table 6. Although some correlations approached significance (learning speed and N1 latency in quiet; learning speed and P3a latency in babble), only a few significant relationships between the physiologic and behavioral measures existed. Specifically, P2 latencies were positively correlated with WR score in quiet, indicating that listeners with longer P2 latencies in quiet performed better on the word recognition task in quiet. P3a latency in babble was also significantly positively correlated with word recognition in babble as well as with learned-word detection, indicating better word recognition and detection of learned words in babble for listeners with longer P3a latencies.

That CAEP component latencies were positively correlated with behavioral performance was an unexpected result. As delayed auditory evoked response latencies are seen in the presence of hearing loss, longer evoked response latencies are generally considered indicative of poorer auditory processing. Thus, it would be expected that superior performance on auditory tasks would be observed in individuals with shorter evoked response latencies, yielding a significant negative correlation between CAEP response latencies and behavioral performance. Given the lack of a consistent pattern of significant correlations between CAEPs and behavioral perforamnce, it is possible that the behavioral tasks used in the present study reflect auditory skills that are unrelated to the CAEPs measured.

CHAPTER 7

DISCUSSION

The objective of this study was to investigate the impact of auditory experience on learning and learning-related mechanics and pragmatics. Individuals with and without long-term musical experience were assessed on their ability to recognize familiar words and to detect, learn, and recall unfamiliar words. CAEPs were measured to examine the impact of enriched auditory experience on cortical functioning reflective of learning. It was hypothesized that participants with extensive musical training experience would demonstrate an advantage over their non-musician counterparts on both electrophysiologic and behavioral measures, and that group differences would be larger for more difficult listening conditions. Neither of these hypotheses was supported by the results in the present study.

Group differences in auditory evoked potentials were consistent with some studies and inconsistent with others. For example, Zhang and colleagues measured CAEPs using monotone pulses and observed longer P2 latencies in musicians than in non-musicians, consistent with the results of the present study (Zhang, Peng, Chen & Hu, 2015). However, a study by Bidelman and Alain found shorter P2 latencies in musicians compared to non-musicians using the same evoking vowel stimuli used in the present study (Bidelman & Alain, 2015). Liang et al. (2016) found similar results using tonal stimuli. Inconsistent findings on evoked potential measures across studies may be an indication that the relationship between music training and cortical auditory processing is either more complex or that these electrophysiologic measures are less sensitive to musical experience than previously thought.

Bidelman and Alain (2015) also found larger P3 amplitudes in musicians compared to non-musicians, whereas P3 amplitude did not differ across group in the present study. A possible explanation for these differing results lies in the specifics of the P300 paradigm used in each study. Participants in Bidelman and Alain (2015) categorized each vowel token presented (as either /u/ or /a/) during CAEP recording. P3 responses, therefore, were measured during an active task that required the listener to make a decision about each stimulus. Thus, the evoked potentials in the Bidelman and Alain study included cortical activity associated with active listening and categorization of auditory stimuli. In the present study, participants completed the categorical perception task separately from CAEP recording, meaning that P3 responses were measured during passive listening. It is possible that including an active task during CAEP measurement changes the response, and that the response is affected differently for individuals with and without music training.

Additionally, the present study differed from the previous work in the presentation of vowel stimuli. In the Bidelman and Alain study, all five vowel tokens were randomly presented in equal proportions (20%). This aspect of the study protocol could not be replicated here, due to limitations of the evoked potential equipment. Instead, stimuli were divided into two sets as described above (i.e., wide and narrow). Within a given set, participants heard random presentations of three of the five vowels, with the vowel with the highest F1 frequency being presented three times as often as either of the other two vowels in the set (see Table 3). Seppänen and colleagues (2012) investigated rapid neural plasticity in young adult musicians and non-musicians by alternating passive and active blocks of tones presented in an oddball paradigm. They observed a decrease in P2 source activation in musicians between the first and second passive blocks. This decrease recovered after musicians resumed an active auditory discrimination task (Seppänen, Hämäläinen, Pesonen, & Tervaniemi, 2012). It may be that weighting stimulus presentation towards higher-pitched vowels caused the listeners to habituate to these sounds to a greater degree than the more rarely presented lower-pitched vowels (recall that responses were collapsed across token type [standard, deviant 1, deviant 2] for the analyses). Habituation, combined with the use of a passive listening task, may have eliminated the musical advantage seen in other studies. While these alterations to the P300 protocol do offer a possible explanation for the lack of significance in the present study, the results also suggest that the advantages of music training may not be particularly substantial if they can be eliminated by small changes to the test protocol.

Turning to the behavioral tasks, the absence of a significant effect of group on word recognition was somewhat surprising, given findings from previous studies showing superior speech-in-noise perception in musicians (Parbery-Clark et al., 2009b; Strait et al., 2012). Thus, it may be useful to consider other aspects of the present study that differed from these prior investigations. In previous studies, speech perception in noise was assessed using tests with varying SNRs, in which either the noise was presented at a constant level and the level of the speech was increased and decreased according to the listener's performance (Hearing-In-Noise Test [HINT; Nilsson, Soli, & Sullivan, 1994]), or the level of the speech was held constant and the level of the noise was gradually increased (Quick Speech-In-Noise test [QuickSIN; Killion, Niquette, Gudmundsen, Revit, & Banerjee, 2004]; Words in Noise test [WIN; Wilson, 1993]).

This allowed speech recognition to be tested at SNRs ranging from 0 to 25 dB for QuickSIN, and < 0 dB for HINT. Performance on these tests was assessed in terms of SNR – either criterion performance (50% of words repeated accurately) or SNR loss (increase in SNR needed to understand speech in noise, relative to normal performance) – rather than percent correct, as was the case in the present study. Results of these previous studies revealed that musicians were able to repeat words and sentences accurately at poorer SNRs than non-musicians. It is possible that, although word recognition performance in the present study was equally impacted by babble (+3 dB SNR) across groups, musicians may have outperformed non-musicians at poorer SNRs. That is, it may be the case that varying levels of noise/babble affect speech recognition differentially across musician and non-musician groups.

Further, the musically-trained children in the study by Strait et al. (2012) showed superior performance on speech-in-noise testing when the noise was spatially separated from the speech signal, but not when noise was collocated with speech. On the other hand, adult musicians outperformed their non-musician peers in collocated but not spatially separated noise conditions (Parbery-Clark et al., 2009b). In the present study, speech and babble were presented binaurally through insert earphones, meaning both ears received identical signals, similar to the collocated noise condition used in previous studies. Thus, the results of word recognition in babble in the present study are inconsistent with previous findings in adult musicians and non-musicians in the study by Parbery-Clark and colleagues. However, Ruggles and colleagues assessed speech understanding in noise in musicians and non-musicians with the goal of replicating the Parbery-Clark study (Ruggles, Freyman, & Oxenham, 2014). While musicians in this study performed better on HINT and QuickSIN than their non-musician peers, group differences were not significant. Fuller et al. tested word recognition at several different SNRs (+10, +5, 0), as well as sentence recognition using several different types of maskers (speech-shaped steady noise, speech-shaped fluctuating noise, 6-talker babble) in musicians and non-musicians (Fuller, Galvin, Maat, Free, & Başkent, 2014). Musicians' scores on word recognition were significantly better than those of non-musicians for only one of the three SNRs (+5 dB), while no significant group differences were observed for sentence recognition in any condition. Similarly, Boebinger and colleagues used four different masker types (clear speech, spectrally rotated speech, speech-amplitude modulated noise, speech-spectrum steady-state noise) to assess sentence recognition in noise in musicians and non-musicians and found no advantage of music training in perceiving masked speech (Boebinger et al., 2015). Taken together, these findings indicate that the advantage of music training for understanding speech in noise, which was lacking in the present study, is perhaps not as robust as initial findings seemed to suggest.

Tests of pragmatics (i.e., learned word retention and detection) assessed participants' ability to draw from information learned the previous day. As discussed above, it has been suggested that music students create internal musical representations by drawing from what they know from previously encountered material when examining a new piece of music (Bamberger, 2000, as cited in Taetle & Cutietta, 2002). The learned word retention and detection tasks used in the present study required participants to access information learned the previous day and identify it in a different modality (retention) or different context (detection). However, musicians, who were presumed to have more experience in using prior knowledge when processing new information compared to non-musicians, did not perform any better on these tasks than the listeners without such extensive experience. That is, skills related to music learning did not appear to transfer to word learning. It is possible that this task is simply not sensitive to differences in long-term auditory experience, such as musical training. However, given the lack of group and condition differences for the Day 1 word-learning and Day 2 learned-word detection tasks, it is more likely that the ability to learn, recall, and detect new words in context is not affected by musical experience.

Active Learning

Participants in both groups of the present study were recruited mainly from the graduate student population at Arizona State University (ASU). While two non-musicians were ASU faculty or staff and four musicians were recruited from local community symphonies, the remaining 85% of participants in this study were active learners (graduate students). It may be the case that current learning in any discipline is sufficient to produce an advantage for word learning, and that the particular tasks used in the present study do not single out any specific advantage of learning music, if one exists. To date, little work has been done to examine differences in learning ability, working memory, auditory attention, etc., between current students and former (or non-) students. Further research is necessary to determine whether active learning in any form influences performance on the auditory learning tasks used in the present study, and whether it does so to a greater extent than auditory experience.

Noise Exposure during Music Training

One aspect of auditory experience that was not considered in the present study was history of noise exposure. Daily noise dose values for college student musicians have been found to exceed 100% of both National Institute for Occupational Safety and Health (NIOSH) and Occupational Safety and Health Administration (OSHA) criteria (Washnik, Phillips, & Teglas, 2016; Miller, Stewart & Lehman, 2007). Previous research also suggests that daily noise exposure for music majors exceeds that of non-music majors (Hawkins, 2013 [significance not reported]).

The effect of noise exposure on hearing is often not evident until it reaches a point sufficient to produce an audiometric “noise notch” – that is, elevated hearing thresholds at ~3-4 kHz. However, the phenomenon of “hidden hearing loss” has gained attention in recent years. It has been suggested that noise exposure producing temporary threshold shifts may cause long-term perceptual difficulties, even after hearing sensitivity for low-level sounds has recovered (e.g., Plack, Barker, & Prendergast, 2014; Schaette & McAlpine, 2011). Considering that musicians tend to be regularly exposed to higher levels of noise than non-musicians, they are presumably at higher risk for the consequences of hidden hearing loss. Stamper and Johnson (2015) measured distortion product otoacoustic emissions (DPOAEs) and auditory brainstem responses (ABRs) in normal-hearing young adults with varying noise exposure backgrounds. Many participants in their study who reported more extensive histories of noise exposure were individuals recruited from the music departments of local universities. Participants’ annual amount of noise exposure was significantly negatively correlated with the amplitude of wave I ABR to a suprathreshold-level click stimulus. By contrast, no systematic relationship between noise exposure history and DPOAE levels was observed. These results suggest that noise exposure can degrade neural function even in the absence of permanent hearing loss (Stamper & Johnson, 2015).

Although no audiometric differences between groups were found in the present study, participants in the musician group likely had more extensive histories of noise exposure compared to the non-musicians with backgrounds in fields such as neuroscience and history. If some of the musicians did indeed have hidden hearing loss, it is possible that its effects counteracted any benefit they may have gleaned from their enriched auditory experience. There may then be a trade-off between the cognitive benefits of music training and the physiologic effects of its associated exposure to loud sounds. Future studies combining the measures used in the present study with assessments of noise exposure and hidden hearing loss may be useful in examining the benefits of music training for auditory tasks, weighted against the costs of noise exposure.

Interpretation of the Results in the Context of the Baltes Model

It is important to note that the participants in this study were representative of the portion of the Baltes model where mechanics, learning, and pragmatics are at their maximum and are expected to differ the least across groups (see Figure 1). Therefore, robust relationships among outcome measures are not as likely to be observed in young normal-hearing listeners who show healthy neural function and generally good performance on listening tasks. The relationship between cognitive mechanics and learning/pragmatic ability may be clearer in a group of listeners with varied physiology (e.g., normal-hearing vs. hearing-impaired) as well as varied auditory experience over the life span. It is possible that musical training did not provide a benefit for performance on word learning tasks because no detriment existed that required compensation in these young, normal-hearing adults. Listeners in this study were not impacted by either developing or declining mechanics, pragmatics, or learning. Thus, these results may serve as a benchmark against which performance in children and older adults with and without musical experience can be compared. Such studies may better reveal the impact of auditory training on mechanics, pragmatics, and learning during vulnerable stages of development or aging.

CHAPTER 8

CONCLUSIONS

In the present study, enriched auditory experience in the form of long-term music training in young, normal-hearing adults neither significantly enhanced nor degraded their ability to detect, learn, or retain new words. Characteristics of physiologic processing (latencies and amplitudes of cortical auditory evoked responses) were also not correlated with performance on learning and pragmatic tasks. These results indicate that new word detection, learning, and retention as well as cortical electophysiologic processing are equivalent in musicians and non-musicians during the young-adult years. The results lay the foundation for the evaluation of a potential musical advantage with advancing age as predicted in the Baltes model of intelligence.

Table 1. Demographic information, audiometric thresholds for octave frequencies 0.25-8 kHz, and music training history for Musician (MUS) and Non-Musician (NOM) groups.

	M:F	Avg. Age		Average Threshold (dB HL)						Music Training
Group	Ratio	(yrs) (SD)	Ear	.25	.5	1	2	4	8	(yrs) (SD)
MUS	12:8	27.0 (3.8)	R	9	9	8	10	8	5	18.8 (4.0)
			L	11	9	8	9	8	3	
NOM	4:16	26.5 (4.5)	R	10	10	9	8	7	5	0.5 (0.7)
			L	11	10	8	8	8	3	

Table 2. Frequencies (Hz) of the fundamental, 1st, 2nd, and 3rd formants of synthetic vowel sounds used to evoke cortical auditory evoked responses (CAEPs).

Vowel	F0	F1	F2	F3
1	100 Hz	430 Hz	1090 Hz	2350 Hz
2	100 Hz	490 Hz	1090 Hz	2350 Hz
3	100 Hz	585 Hz	1090 Hz	2350 Hz
4	100 Hz	660 Hz	1090 Hz	2350 Hz
5	100 Hz	730 Hz	1090 Hz	2350 Hz

Table 3. Vowel token distribution for wide and narrow stimulus sets used in cortical auditory evoked potential (CAEP) testing.

		Wide		Narrow	
	%	Vowel	F1	Vowel	F1
Standard	60	5	730 Hz	4	660 Hz
Deviant 1	20	3	585 Hz	3	585 Hz
Deviant 2	20	1	430 Hz	2	490 Hz

Table 4. Results of repeated measures analysis of variance for cortical auditory evoked potentials (CAEPs). Bold text indicates significant differences.

	Group		Condition		Group*Condition	
	F	*p*	*F*	*p*	*F*	*p*
N1 Latency	0.13	0.721	**6.43**	**0.015**	0.01	0.926
P2 Latency	**7.55**	**0.003**	**14.68**	**<0.001**	0.05	0.821
N1-P2 Amplitude	4.01	0.052	**18.96**	**<0.001**	1.97	0.169
P3a Latency	0.54	0.469	**4.16**	**0.049**	1.43	0.239
P3a Amplitude	0.43	0.517	**8.12**	**0.007**	4.01	0.052

Table 5. Results of repeated measures analysis of variance for cortical auditory evoked potentials (CAEPs) across listening condition (quiet, babble) for Musicians (MUS) and Non-Musicians (NOM). Bold text indicates significant differences.

	MUS		NOM	
	F	*p*	*F*	*p*
N1 Latency	2.74	0.114	3.79	0.067
P2 Latency	**8.93**	**0.008**	**6.02**	**0.024**
N1-P2 Amplitude	**7.23**	**0.015**	**11.86**	**0.003**
P3a Latency	3.75	0.068	0.59	0.453
P3a Amplitude	0.36	0.557	**11.79**	**0.003**

Table 6. Results of bivariate correlational analyses for electrophysiologic and behavioral measures. Bold text indicates significant correlations.

Quiet

	N1 Lat		P2 Lat		N1-P2 Amp		P3a Lat		P3a Amp	
	r	*p*	*r*	*p*	*r*	*p*	*r*	*p*	*r*	*p*
WR (arcsin)	0.05	0.760	**0.32**	**0.046**	0.24	0.132	-0.09	0.593	0.13	0.431
NWD *d*-prime	-0.15	0.361	-0.23	0.147	0.09	0.601	-0.14	0.407	-0.05	0.780
WL Speed	0.31	0.056	0.20	0.209	-0.08	0.647	-0.10	0.561	0.28	0.082
Retention (arcsin)	0.11	0.483	0.17	0.283	0.04	0.788	-0.01	0.977	0.25	0.117
LWD *d*-prime	0.09	0.590	-0.05	0.778	-0.11	0.504	0.21	0.200	0.01	0.962

Babble

	N1 Lat		P2 Lat		N1-P2 Amp		P3a Lat		P3a Amp	
	r	*p*	*r*	*p*	*r*	*p*	*r*	*p*	*r*	*p*
WR (arcsin)	0.15	0.343	-0.24	0.130	-0.16	0.322	**0.36**	**0.022**	0.02	0.922
NWD *d*-prime	0.09	0.571	0.05	0.782	-0.04	0.815	-0.11	0.501	0.08	0.613
WL Speed	0.06	0.700	-0.04	0.794	0.02	0.891	0.30	0.064	-0.08	0.630
Retention (arcsin)	-0.03	0.875	-0.29	0.071	-0.14	0.384	0.03	0.871	<0.01	0.981
LWD *d*-prime	0.16	0.323	-0.21	0.211	-0.15	0.377	**0.43**	**0.007**	-0.12	0.454

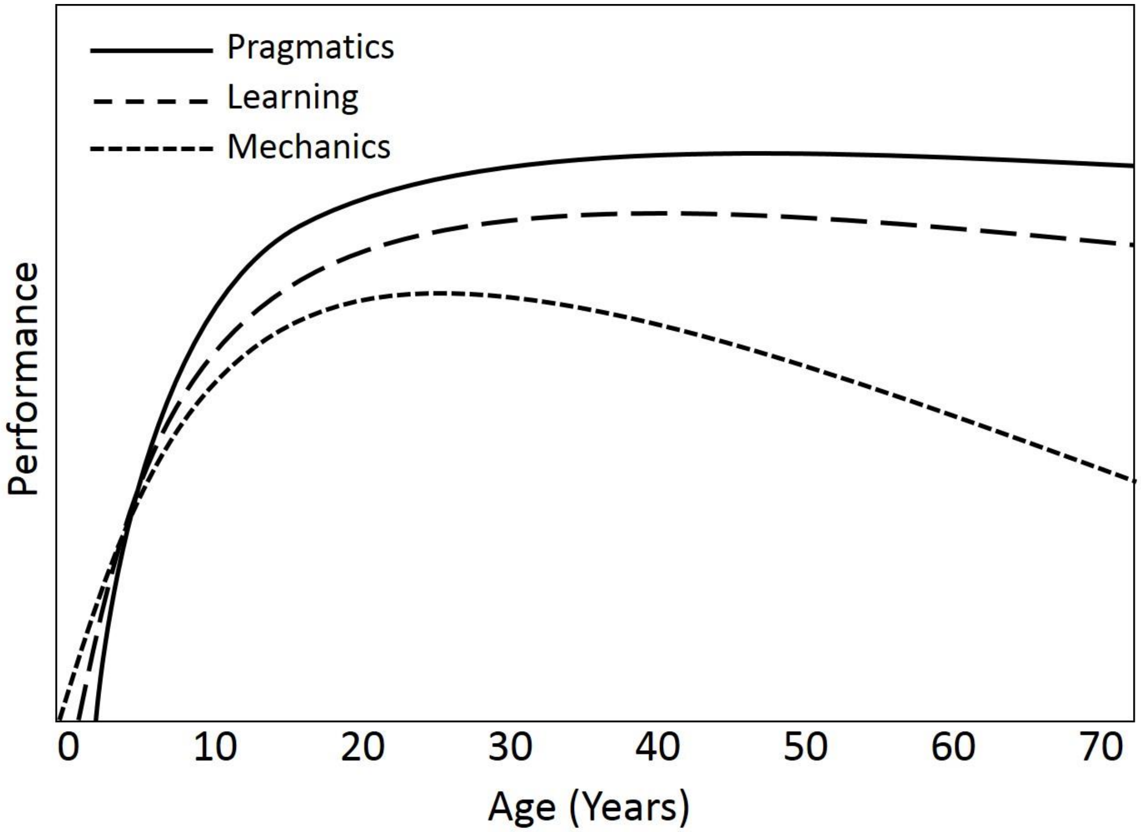


Figure 1. Proposed expansion of the Baltes dual-process model of intelligence (cognitive mechanics, cognitive pragmatics).

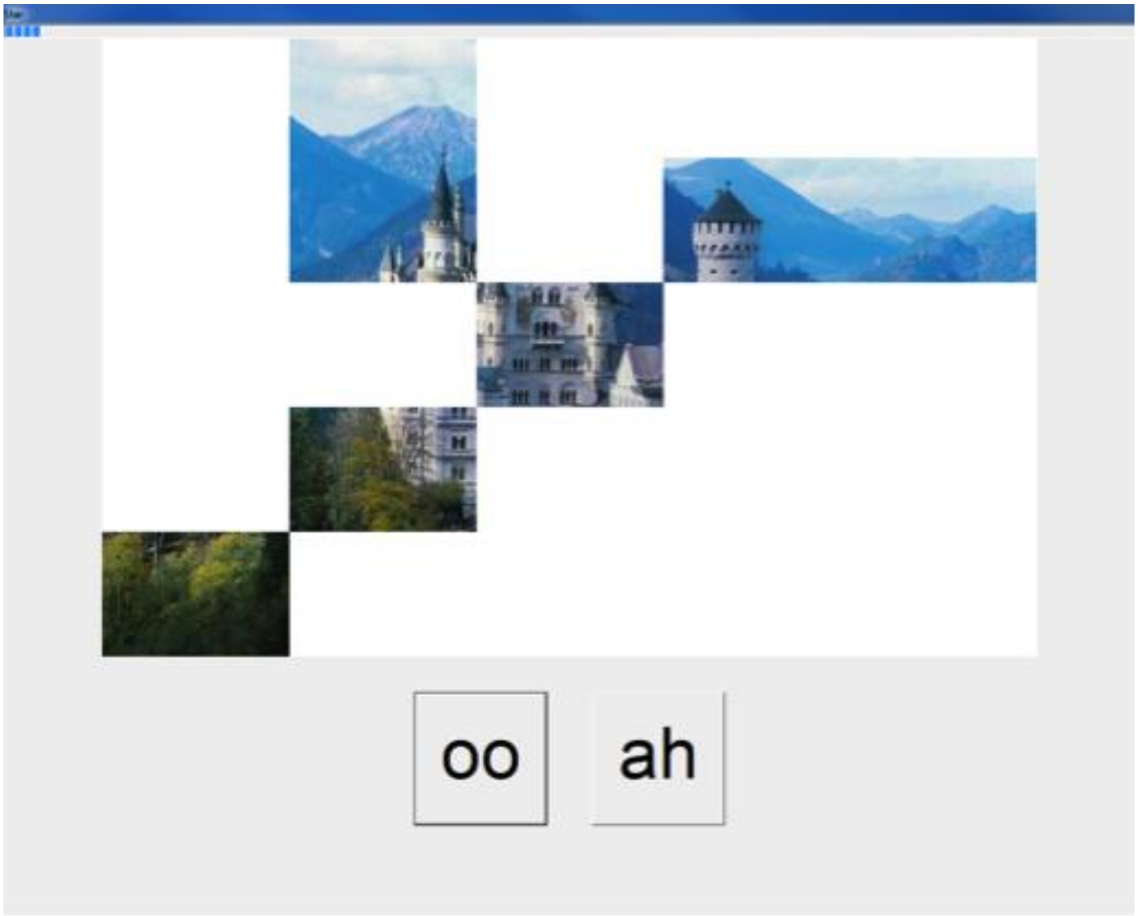


Figure 2. Computer interface used during the categorical perception task.

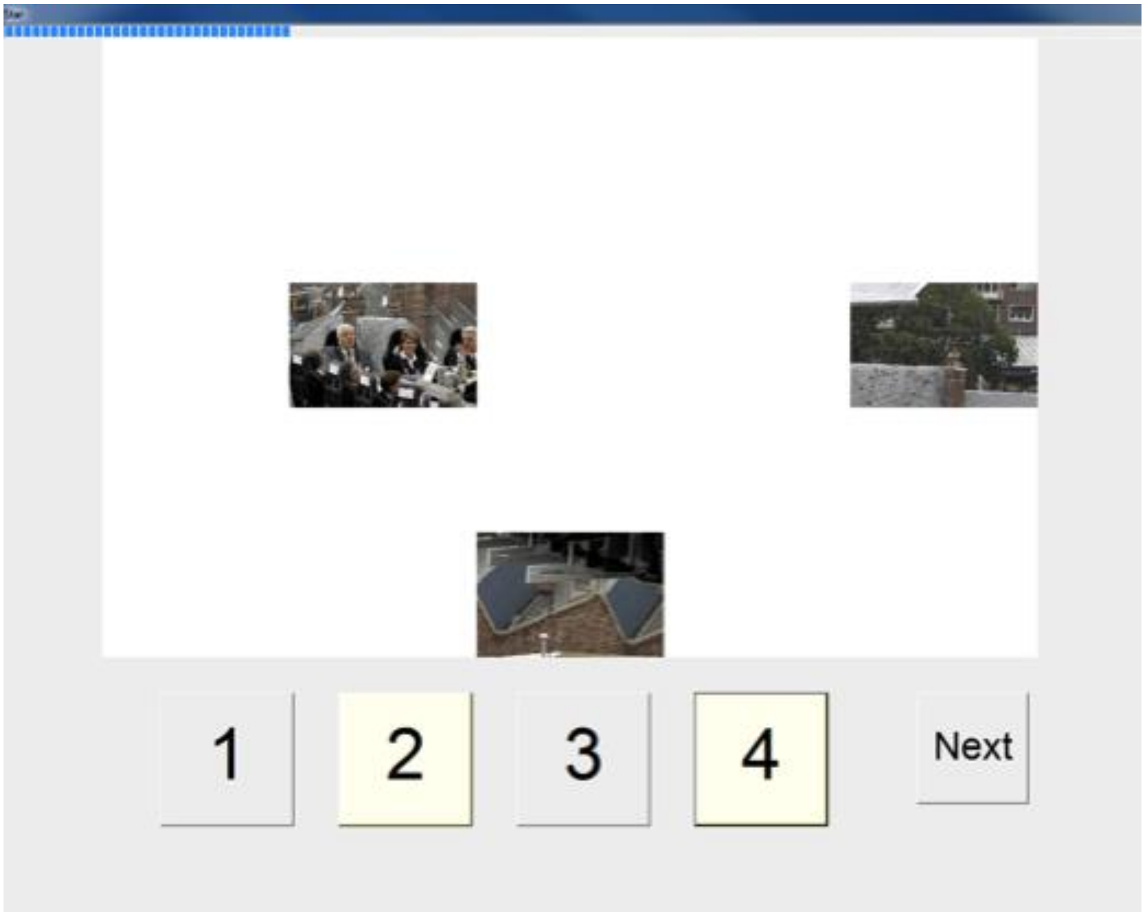


Figure 3. Computer interface used during the non-word detection task. Highlighted buttons indicate the words in the sentence the listener identified/selected as nonsense words.

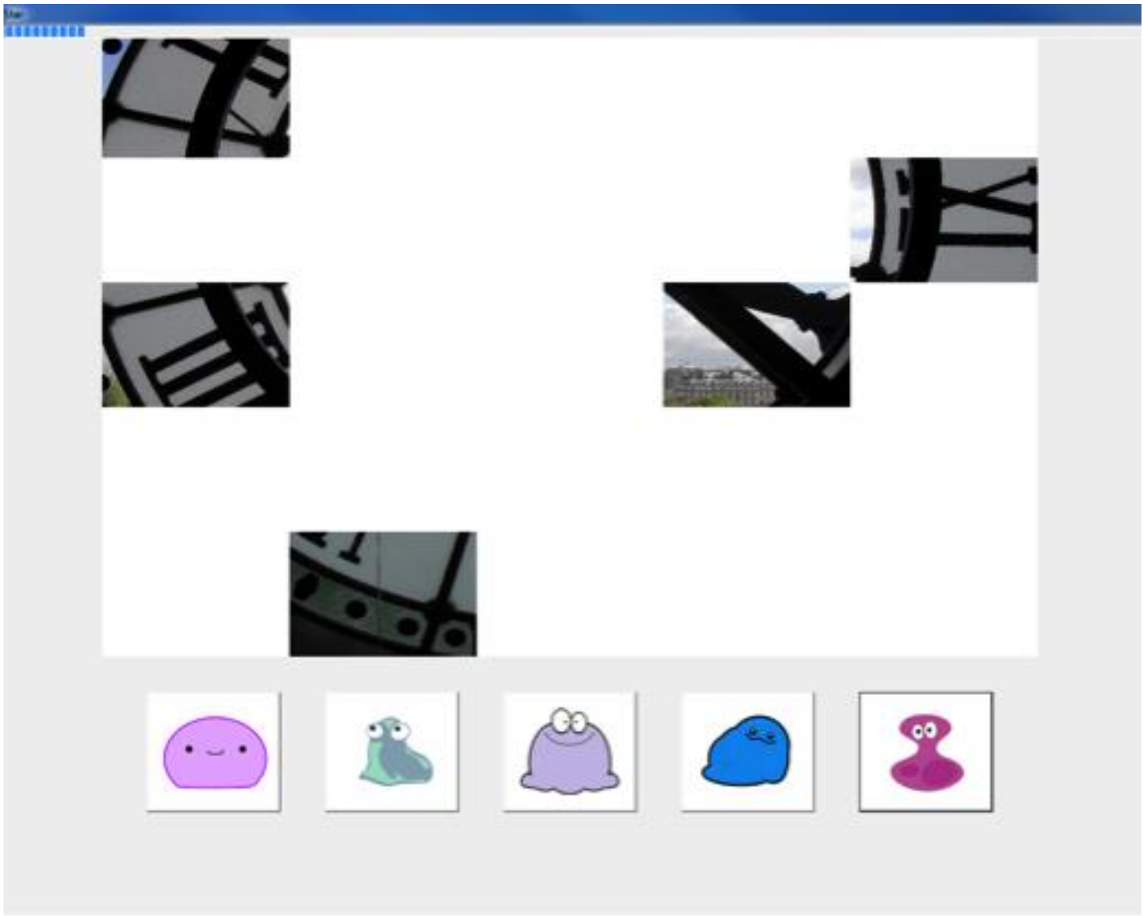


Figure 4. Computer interface used during the word-learning task.

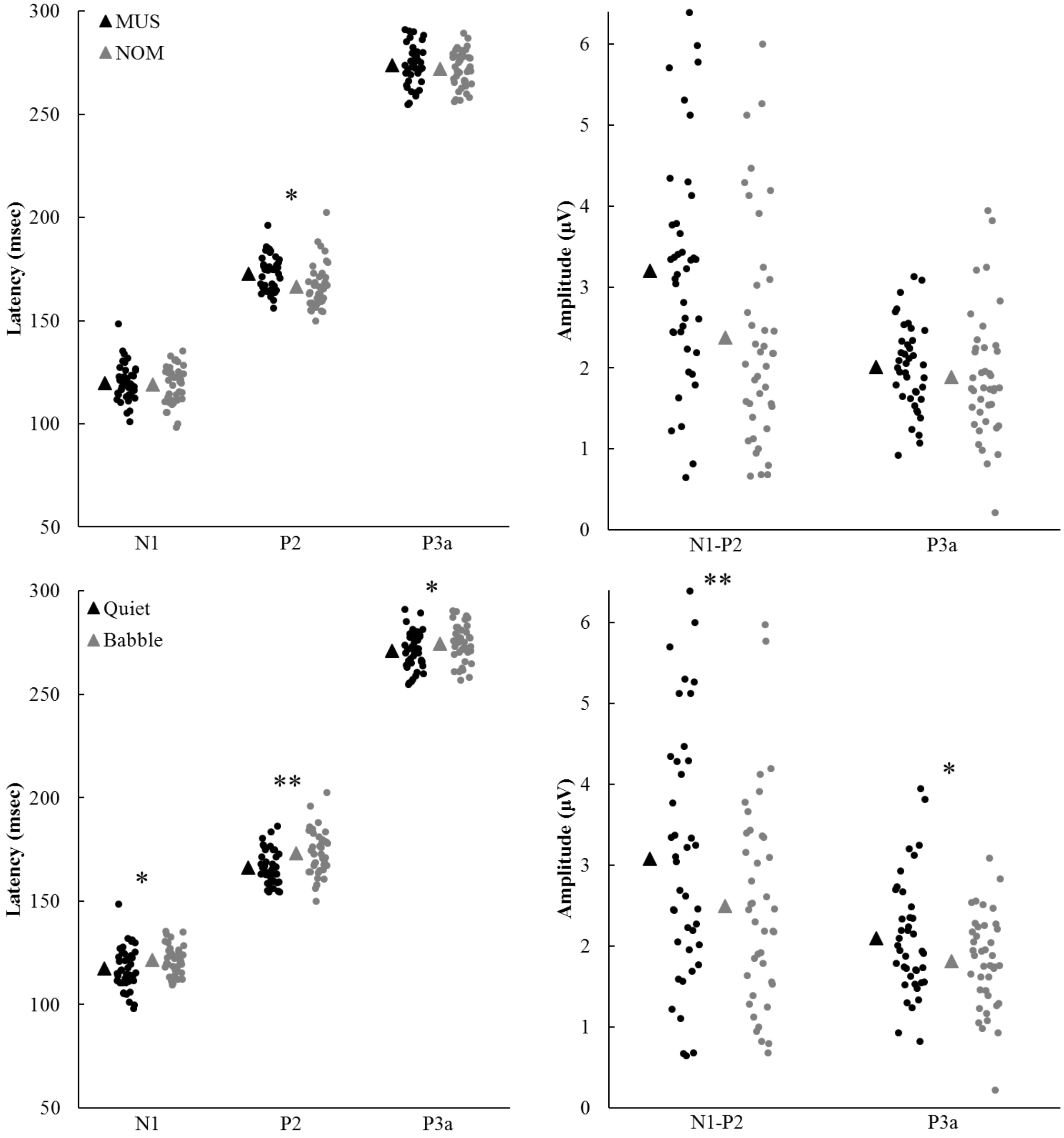


Figure 5. Individual (circles) and average (triangles) latencies and amplitudes of cortical auditory evoked potentials across group (top panels) and listening condition (bottom panels). Asterisks indicate significance (*p<0.05; **p<0.001).

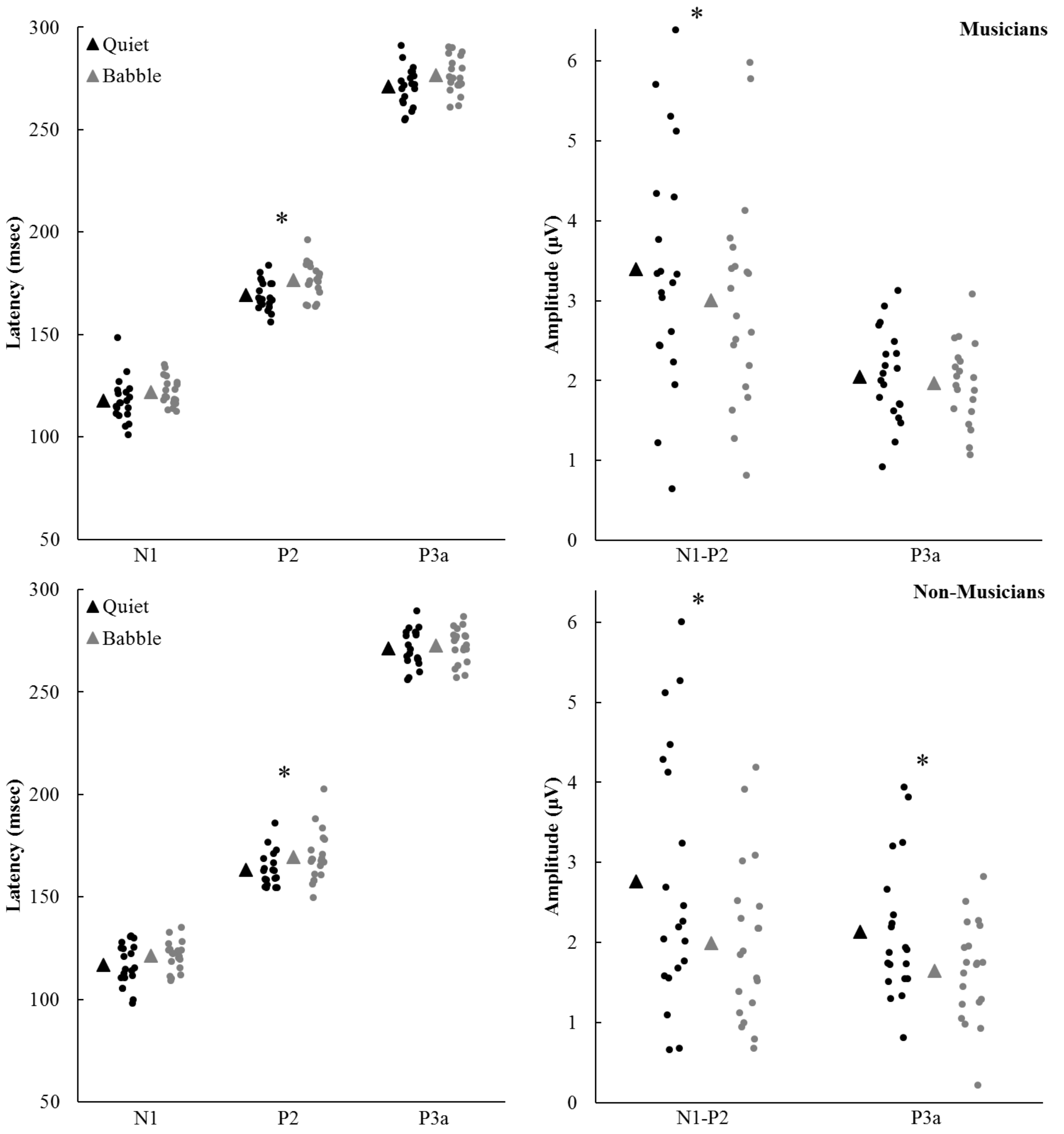


Figure 6. Individual (circles) and average (triangles) latencies and amplitudes of cortical auditory evoked potentials across condition for musicians (top panels) and non-musicians (bottom panels). Asterisks indicate significance (*p<0.05).

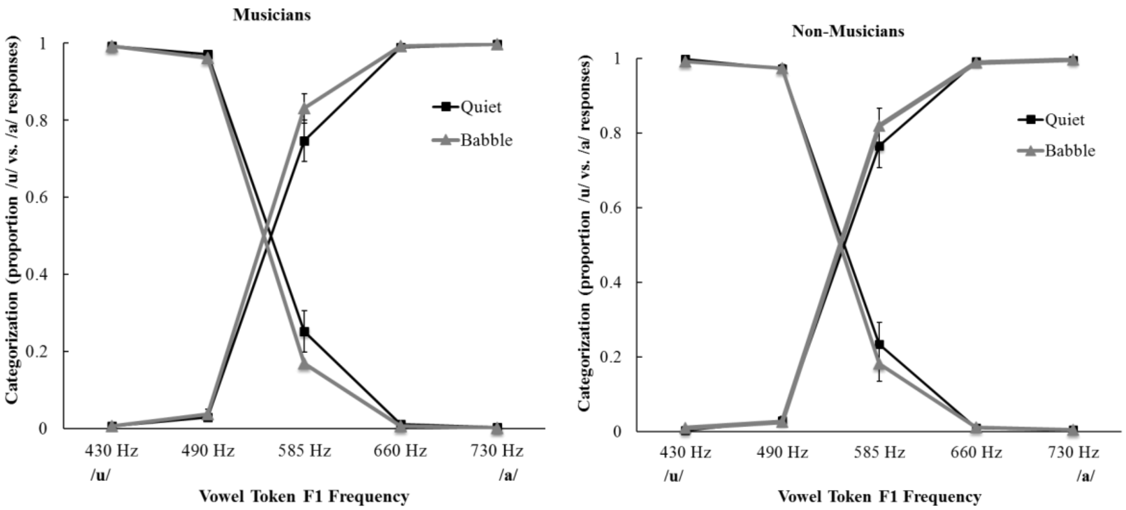


Figure 7. Psychometric functions for /u/ vs. /a/ in quiet and in babble for musicians (left panel) and non-musicians (right panel).

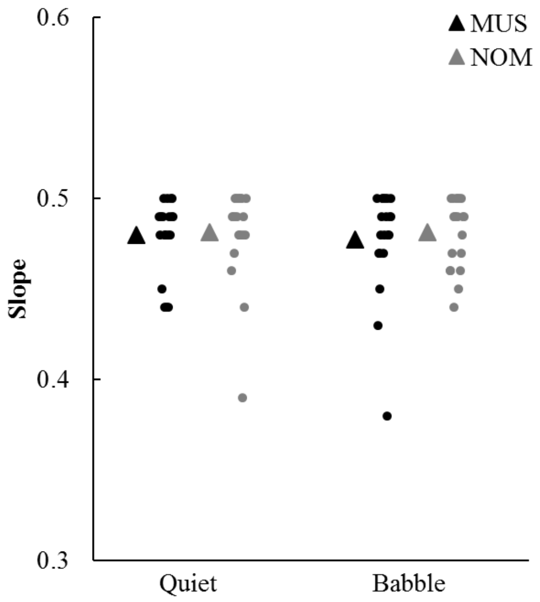


Figure 8. Individual (circles) and average (triangles) slopes (defined in Chapter 5) of the psychometric functions for /u/ vs. /a/ in quiet and in babble for musicians and non-musicians.

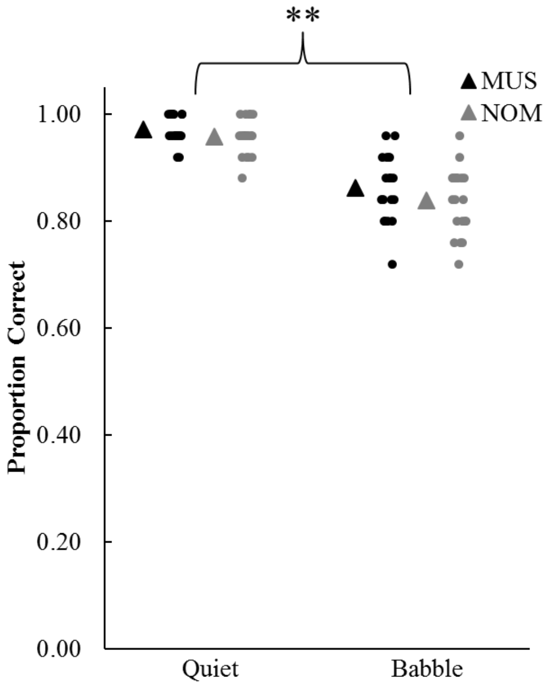


Figure 9. Individual (circles) and average (triangles) scores on word recognition in quiet and babble for musicians (MUS) and non-musicians (NOM). Asterisks indicate significant differences across condition (***p*<0.001).

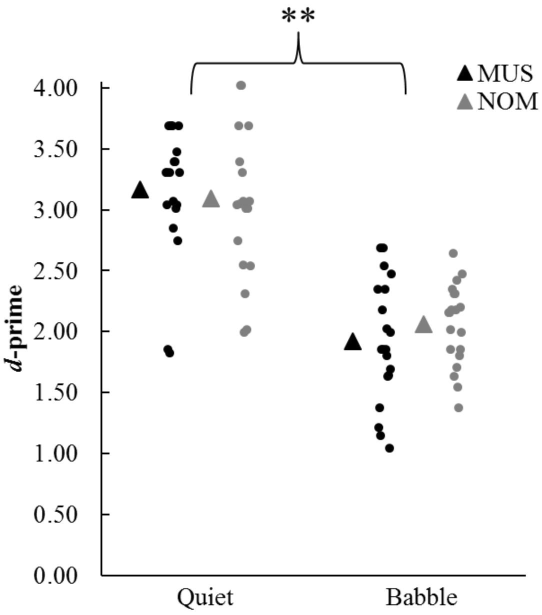


Figure 10. Individual (circles) and average (triangles) scores on non-word detection in quiet and babble for musicians (MUS) and non-musicians (NOM). Asterisks indicate significant differences across condition (***p*<0.001).

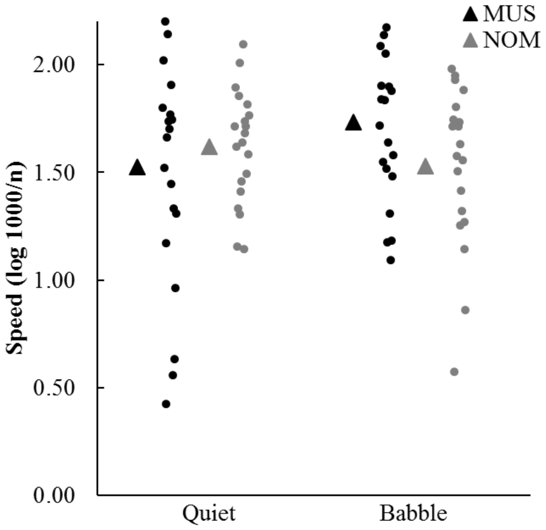


Figure 11. Individual (circles) and average (triangles) speed of word learning in quiet and babble for musicians (MUS) and non-musicians (NOM).

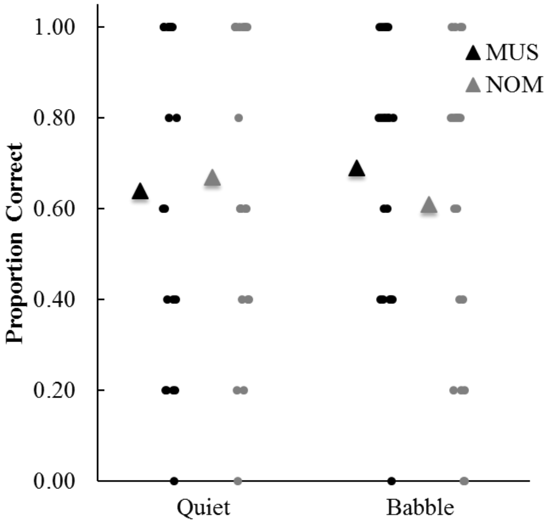


Figure 12. Individual (circles) and average (triangles) scores on next-day word retention in quiet and babble for musicians (MUS) and non-musicians (NOM).

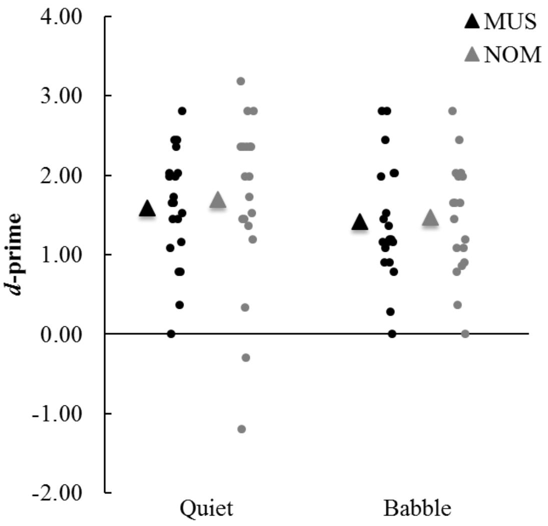


Figure 13. Individual (circles) and average (triangles) scores on learned-word detection in quiet and babble for musicians (MUS) and non-musicians (NOM).

REFERENCES

Alloway, T. P. (2007). *Automated Working Memory Assessment*. London: Psychological Corporation.

Axelsson, E.L., Churchley, K., & Horst, J.S. (2012). The right thing at the right time: why ostensive naming facilitates word learning. *Frontiers in Psychology, 3*, 1-8.

Ayotte, J., Peretz, I., & Hyde, K. (2002). Congenital amusia: A group study of adults afflicted with a music-specific disorder. *Brain, 125*(2), 238-251.

Baddeley, A. D. (1992). Working memory. *Science, 255*(5044), 556-559.

Bahrick, L. E., and Lickliter, R. (2000). Intersensory redundancy guides attentional selectivity and perceptual learning in infancy. *Developmental Psychology, 36*, 190-201.

Bahrick, L. E., Lickliter, R., and Flom, R. (2003). Up versus down: The role of intersensory redundancy in the development of infants' sensitivity to the orientation of moving objects. *Infancy, 9,* 73-96.

Bahrick, L. E., Lickliter, R., and Flom, R. (2004). Intersensory redundancy guides the development of selective attention, perception, and cognition in infancy. *Current Directions in Psychological Science, 13,* 99-102.

Baltes, P. B. (1993). The aging mind: Potential and limits. *The Gerontologist, 33*(5), 580-594.

Beals, D.E. (1997). Sources of support for learning words in conversation. *Journal of Child Language, 24*(3), 673-694.

Besson, M., Chobert, J., & Marie, C. (2011). Transfer of training between music and speech: common processing, attention, and memory. *Frontiers in Psychology,* 2:94.

Bialystok, E., Craik, F., & Luk, G. (2008). Cognitive control and lexical access in younger and older bilinguals. Journal of Experimental Psychology: Learning, *Memory, and Cognition, 34*(4), 859-873.

Bidelman, G. M., & Alain, C. (2015). Musical training orchestrates coordinated neuroplasticity in auditory brainstem and cortex to counteract age-related declines in categorical vowel perception. *The Journal of Neuroscience*, *35*(3), 1240-1249.

Bidelman, G. M., Villafuerte, J. W., Moreno, S., & Alain, C. (2014a). Age-related changes in the subcortical-cortical encoding and categorical perception of speech. *Neurobiology of Aging*, *35*(11), 2526-2540.

Bidelman, G. M., Weiss, M. W., Moreno, S., & Alain, C. (2014b). Coordinated plasticity in brainstem and auditory cortex contributes to enhanced categorical speech perception in musicians. *European Journal of Neuroscience, 40*(4), 2662-2673.

Bloom, P. (2000). *How children learn the meanings of words*. Cambridge, MA: MIT Press.

Blamey, P. J., Sarant, J. Z., Paatsch, L. E., Barry, J. G., Bow, C. P., Wales, R. J., …Tooher, R. (2001). Relationships among speech perception, production, language, hearing loss, and age in children with impaired hearing. *Journal of Speech, Language, and Hearing Research, 44*(2), 264-285.

Boebinger, D., Evans, S., Rosen, S., Lima, C. F., Manly, T., & Scott, S. K. (2015). Musicians and non-musicians are equally adept at perceiving masked speech. *The Journal of the Acoustical Society of America*, *137*(1), 378-387.

Booth, G.D., & Cutietta, R.A. (1991). The applicability of verbal processing strategies to recall of familiar songs. *Journal of Research in Music Education, 39*(2), 121-131.

Brown, R. (1957). Linguistic determinism and the part of speech. *Journal of Abnormal and Social Psychology, 55,* 1-5.

Brown, C. A., & Bacon, S. P. (2010). Fundamental frequency and speech intelligibility in background noise. *Hearing Research*, *266*(1), 52-59.

Brown, J. I., Bennett, J. M., & Hanna, G. (1981). Nelson-Denny Reading Test. Chicago: Riverside.

Bugos, J.A., Perlstein, W.M., McCrae, C.S., Brophy, T.S., & Bedenbaugh, P.H. (2007). Individualized piano instruction enhances executive functioning and working memory in older adults. *Aging & Mental Health*, *11*(4), 464-471.

Capone, N.C., & McGregor, K.K. (2005). The effect of semantic representation on toddlers' word retrieval. *Journal of Speech, Language, and Hearing Research, 48*(6), 1468-1480.

Carey, S., & Bartlett, E. (1978). Acquiring a single new word. *Paper and Reports on Child Language Development*, *15*, 17-29.

Caspary, D. M., Ling, L., Turner, J. G., & Hughes, L. F. (2008). Inhibitory neurotransmission, plasticity and aging in the mammalian central auditory system. *The Journal of Experimental Biology*, *211*(11), 1781-1791.

Chandrasekaran B., & Kraus, N. (2010). Music, noise-exclusion, and learning. *Music Perception*, *27*(4), 297-306.

Clark, T., Williamon, A., & Aksentijevic, A. (2012). Musical imagery and imagination: The function, measurement, and application of imagery skills for performance. In Hargreaves, D., Miell, D., & MacDonald, R. (Eds.) *Musical imaginations: Multidisciplinary perspectives on creativity, performance, and perception.* Oxford: Oxford University Press.

Corrigall, K. A., & Trainor, L. J. (2011). Associations between length of music training and reading skills in children. *Music Perception*, *29*(2), 147-155.

Cole, M. (1996). *Cultural psychology: A once and future discipline*. Cambridge, MA: Belknap Press of Harvard University Press.

Daneman, M., & Carpenter, P. A. (1983). Individual differences in integrating information between and within sentences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *9*(4), 561-584.

Dunn, L. M., & Dunn, L. M. (1981). Peabody Picture Vocabulary Test – Revised. Circle Pines, MN: American Guidance Service.

Fuller, C. D., Galvin III, J. J., Maat, B., Free, R. H., & Başkent, D. (2014). The musician effect: does it persist under degraded pitch conditions of cochlear implant simulations? *Frontiers in Neuroscience*, 8:179.

Gordon, E. (2003). *Learning sequences in music: Skill, content, and patterns; a music learning theory* (2003 ed.). Chicago, Ill.: GIA Publications.

Goy, H., Pichora-Fuller, M. K., Singh, G., Russo, F. A. (2016). Perception of emotional speech by listeners with hearing aids. *Journal of the Canadian Acoustical Association*, *44*(3).

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

Grose, J. H., Mamo, S. K., & Hall, J. W. (2009). Age effects in temporal envelope processing: speech unmasking and auditory steady state responses. *Ear & Hearing*, *30*(5), 568-575

Hanna-Pladdy, B., & MacKay, A. (2011). The relation between instrumental musical activity and cognitive aging. *Neuropsychology*, *25*(3), 378-386.

Hannon, B., & Daneman, M. (2009). Age-related changes in reading comprehension: an individual-differences perspective. *Experimental Aging Research*, *35*(4), 432-456.

Hargreaves, D.J., & Zimmerman, M.P. (1992). Developmental theories of music learning. In Colwell, R. (Ed.) *Handbook of Research on Music Teaching and Learning: A project of the Music Educators National Conference*. New York: Schirmer Books.

Hawkins, E. F. (2013). "Comparison of hearing levels of college music and non-music majors: Does rehearsal noise affect hearing health?" Independent Studies and Capstones. Paper 678. Program in Audiology and Communication Sciences, Washington University School of Medicine.

Heibeck, T.H., & Markman, E.M. (1987). Word learning in children: an examination of fast mapping. *Child Development*, *58*(4), 1021-1034.

Helfer, K. S., & Wilber, L. A. (1990). Hearing loss, aging, and speech perception in reverberation and noise. *Journal of Speech and Hearing Research*, *33*(1), 149-155.

Horst, J.S., & Samuelson, L.K. (2008). Fast mapping but poor retention by 24-month-old infants. *Infancy*, *13*(2), 128-157.

Jonides, J., Lewis, R. L., Nee, D. E., Lustig, C. A., Berman, M. G., & Moore, K. S. (2008). The mind and brain of short-term memory. *Annual Review of Psychology, 59*, 193-224.

Killion, M. C., Niquette, P. A., Gudmundsen, G. I., Revit, L. J., & Banerjee, S. (2004). Development of a quick speech-in-noise test for measuring signal-to-noise ratio loss in normal-hearing and hearing-impaired listeners. *Journal of the Acoustical Society of America, 116* (4 Pt 1), 2395-2405.

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

Kratus, J. (1994). Relationships among children's music audiation and their compositional processes and products. *Journal of Research in Music Education, 42*(2), 115-130.

Kraus, N., Skoe, E., Parbery-Clark, A., & Ashley, R. (2009). Experience-induced malleability in neural encoding of pitch, timbre, and timing. *Annals of the New York Academy of Sciences, 1169*(1), 543-557.

Lee, K. M., Skoe, E., Kraus, N., & Ashley, R. (2009). Selective subcortical enhancement of musical intervals in musicians. *The Journal of Neuroscience, 29*(18), 5832-5840.

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

Lorenzi, C., Gilbert, G., Carn, H., Garnier, S., and Moore, B. C. J. (2006). Speech perception problems of the hearing impaired reflect inability to use temporal fine structure. *Proceedings of the National Academy of Sciences, 103*(49), 18866-18869.

Luria, A. R., Tsvetkova, L. S., & Futer, D. S. (1965). Aphasia in a composer. *Journal of the Neurological Sciences, 2*, 288-292.

McEvoy, L. K., Pellouchoud, E., Smith, M. E., & Gevins, A. (2001). Neurophysiological signals of working memory in normal aging. *Brain Research: Cognitive Brain*

Research, *11*(3), 363-376.

Miller, G.A., & Gildea, P.M. (1987). How children learn words. *Scientific American, 257*(3), 94-99.

Miller, V. L., Stewart, M., & Lehman, M. (2007). Noise exposure levels for student musicians. *Medical Problems of Performing Artists*, *22*(4), 160-165.

Mishra, S. K., Panda, M. R., & Raj, S. (2015). Influence of musical training on sensitivity to temporal fine structure. *International Journal of Audiology*, *54*(4), 220-226.

Moore, B. C. J. (2008). The role of temporal fine structure processing in pitch perception, masking, and speech perception for normal-hearing and hearing-impaired people. *Journal of the Association for Research in Otolaryngology*, *9*(4), 299-406.

Moreno-Gómez, F. N., Véliz, G., Rojas, M., Martínez, C., Olmedo, R., Panussis, F., … Delano, P. H. (2017). Music training and education slow the deterioration of music perception produced by presbycusis in the elderly. *Frontiers in Aging Neuroscience*, 9:149.

Musacchia, G., Sams, M., Skoe, E., & Kraus, N. (2007). Musicians have enhanced subcortical auditory and audiovisual processing of speech and music. *Proceedings of the National Academy of Sciences*, *104*(40), 15894-15898.

Musacchia, G., Strait, D., & Kraus, N. (2008). Relationships between behavior, brainstem and cortical encoding of seen and heard speech in musicians and non-musicians. *Hearing Research*, *241*(1), 34-42.

Nagy, W., Herman, P., & Anderson, R. (1985). Learning words from context. *Reading Research Quarterly*, *20*, 233-253.

Nilsson, M., Soli, S., & Sullivan, J. (1994). Development of the Hearing in Noise Test for the measurement of speech reception thresholds in quiet and in noise. *Journal of the Acoustical Society of America*, *95*(2), 1085-1099.

Nutley, S. B., Darki, F., & Klingber, T. (2014). Music practice is associated with development of working memory during childhood and adolescence. *Frontiers in Human Neuroscience*, 7:926.

Pantev, C., Ross, B., Fujioka, T., Trainor, L. J., Schulte, M., & Schulz M. (2003). Music and learning-induced cortical plasticity. *Annals of the New York Academy of Sciences*, *999*(1), 438-450.

Partanen, E., Kujala, T., Näätänen, R., Liitola, A., Sambeth, A., & Huotilainen, M. (2013). Learning-induced neural plasticity of speech processing before birth. *Proceedings of the National Academy of Sciences*, *110*(37), 15145-15150.

Parbery-Clark, A., Anderson, S., Hittner, E., & Kraus, N. (2012). Musical experience offsets age-related delays in neural timing. *Neurobiology of Aging*, *33*(7), 1483.e1-1483.e4.

Parbery-Clark, A. Skoe, E., & Kraus, N. (2009a). Musical experience limits the degradative effects of background noise on the neural processing of sound. *The Journal of Neuroscience*, *29*(45), 14100-14107.

Parbery-Clark, A., Skoe, E., Lam, C., & Kraus, N. (2009b). Musician enhancement for speech-in-noise. *Ear & Hearing*, *30*(6), 653-61.

Patel, A. D. (2008). Neural resources for syntactic integration as a key link. In *Music, Language, and the Brain* (pp. 267-297). Oxford, NY: Oxford University Press.

Patel, A. D. (2003). Language, music, syntax and the brain. *Nature Neuroscience*, *6*(7), 674-681.

Peretz, I. (1993) Auditory atonalia for melodies. *Cognitive Neuropsychology*, *10*(1), 21-56.

Peretz, I., Champod, A. S., & Hyde, K. (2003). Varieties of musical disorders. The Montreal Battery of Evaluation of Amusia. *Annals of the New York Academy of Sciences*, *999*, 58-75.

Peters, R. W., & Hall, J. W. (1994). Comodulation masking release for elderly listeners with relatively normal audiograms. *The Journal of the Acoustical Society of America*, *96*(5 Pt 1), 2674-2682.

Pinal, D., Zurrón, M., & Díaz, F. (2015). Age-related changes in brain activity are specific for high order cognitive processes during successful encoding of information in working memory. *Frontiers in Aging Neuroscience*, 7:75.

Pittman, A. L., & Rash, M. A. (2016). Auditory lexical decision and repetition in children: Effects of acoustic and lexical constraints. *Ear & Hearing*, *37*(2), 119-128.

Pittman, A. (2011). Age-related benefits of digital noise reduction for short-term word learning in children with hearing loss. *Journal of Speech Language and Hearing Research*, *54*(5), 1448-1463.

Pittman, A. L. (2008). Short-term word-learning rate in children with normal hearing and children with hearing loss in limited and extended high-frequency bandwidths. *Journal of Speech Language and Hearing Research*, *51*(3), 785-797.

Pittman, A. L., Lewis, D. E., Hoover, B. M., & Stelmachowicz, P. G. (2005). Rapid word-learning in normal-hearing and hearing-impaired children: effects of age, receptive vocabulary, and high-frequency amplification. *Ear & Hearing*, *26*(6), 619-629.

Pittman, A. L., Stewart, E. C., Willman, A., & Odgear, I. S. (2017). Word recognition and learning: effects of hearing loss and amplification feature. *Trends in Hearing*, *21*, 1-13.

Plack, C. J., Barker, D., & Prendergast, G. (2014). Perceptual consequences of "hidden" hearing loss. *Trends in Hearing*, *18*, 1-11.

Polich, J., Howard, L., & Starr, A. (1983). P300 latency correlates with digit span. *Psychophysiology*, *20*(6), 665-669.

Purcell, D. W., John, S. M., Schneider, B. A., & Picton, T. W. (2004). Human temporal auditory acuity as assessed by envelope following responses. *The Journal of the Acoustical Society of America*, *166*(6), 3581-3593.

Quine, W. (1960). *Word and object*. Cambridge: Technology Press of the Massachusetts Institute of Technology.

Rice, M. L., Buhr, J. C., & Nemeth, M. (1990). Fast mapping word-learning abilities of language-delayed preschoolers. *Journal of Speech and Hearing Disorders*, *55*(1), 33-42.

Rosen, S. (1992). Temporal information in speech: acoustic, auditory and linguistic aspects. *Philosophical Transactions of the Royal Society B, 336*(1278), 367-373.

Ruggles, D. R., Freyman, R. L., & Oxenham, A. J. (2014). Influence of musical training on understanding voiced and whispered speech in noise. *PLoS One, 9*(1), e86980.

Schaette, R., & McAlpine, D. (2011). Tinnitus with a normal audiogram: physiological evidence for hidden hearing loss and computational model. *The Journal of Neuroscience*, *31*(38), 13452-13457.

Schellenberg, E. G. (2001). Music and nonmusical abilities. *Annals of the New York Academy of Sciences, 930*(1), 355-371.

Schnupp, J., Nelken, I., & King, A. (2011). Auditory Scene Analysis. In *Auditory Neuroscience: Making Sense of Sound*. Cambridge, MA: MIT Press.

Schön, D., Magne, C., & Besson, M. (2004). The music of speech: Music training facilitates pitch processing in both music and language. *Psychophysiology*, *41*(3), 341-349.

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

Seppänen, M., Hämäläinen, J., Pesonen, A., & Tervaniemi, M. (2012). Music training enhances rapid neural plasticity of N1 and P2 source activation for unattended sounds. *Frontiers in Human Neuroscience*, 6:43

Shahin, A., Bosnyak, D. J., Trainor, L. J., & Roberts, L. E. (2003). Enhancement of neuroplastic P2 and N1c auditory evoked potentials in musicians. *The Journal of Neuroscience, 23*(12), 5545-5552.

Sloutsky, V. M., & Yao, X. (2008). Learning words from context. In B. C. Love, K. McRae, & V. M. Sloutsky (Eds.), *Proceedings of the 30th Annual Conference of the Cognitive Science Society* (pp. 1925-1929). Austin, TX: Cognitive Science Society.

Snyder E., & Hillyard, S. A. (1976). Long-latency evoked potentials to irrelevant deviant stimuli. *Behavioral Biology, 16*, 319-331.

Squires, N. K., Squires, K. C., & Hillyard, S. A. (1975). Two varieties of long-latency positive waves evoked by unpredictable auditory stimuli in man. *Electroencephalography and Clinical Neurophysiology*, *38*(4), 387-401.

Stamper G. C., & Tiffany A. Johnson, T. A. (2015). Auditory function in normal-hearing, noise-exposed human ears. *Ear & Hearing*, *36*(2), 172-184.

Stern, Y. (2002). What is cognitive reserve? Theory and research application of the reserve concept. *Journal of the International Neuropsychological Society*, *8*, 448-460.

Stern, Y. (2009). Cognitive reserve. *Neuropsychologia*, *47*, 2015-2028.

Sternberg, R.J. (1987). Most vocabulary is learned from context. In M.G. McKeown & M.E. Curtis (Eds.), The Nature of Vocabulary Acquisition. Hilldale, NJ: Erlbaum.

Sternberg, R. J., & Powell, J. S. (1983). Comprehending verbal comprehension. *American Psychologist*, *38*(8), 878-893.

Strait, D. L., Kraus, N., Parbery-Clark, A., & Ashley, R. (2010). Musical experience shapes top-down mechanisms: Evidence from masking and auditory attention performance. *Hearing Research*, *261*, 22-29.

Strait, D. L., Kraus, N., Skoe, E., & Ashley, R. (2009). Musical experience and neural efficiency - effects of training on subcortical processing of vocal expressions of emotion. *European Journal of Neuroscience*, *29*(3), 661-668.

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

Taetle, L., & Cutietta, R. (2002). Learning Theories as Roots of Current Musical Practice and Research. In R. Colwell & C. Richardson (Eds.), *The New Handbook of Research on Music Teaching and Learning: A Project of the Music Educators National Conference*. Oxford: Oxford University Press.

Takahashi, G. A., & Bacon, S. P. (1992). Modulation detection, modulation masking, and speech understanding in noise in the elderly. *Journal of Speech and Hearing Research*, *35*(6), 1410-1421.

Tillman, T. W., & Carhart, R. (1966). An expanded test for speech discrimination utilizing CNC monosyllabic words. Northwestern University Auditory Test No. 6. SAM-TR-66-55. *Tech. Rep. SAM. -TR*, 1-12.

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

Washnik, N. J., Phillips, S. L., Teglas, S. (2016). Student's music exposure: Full-day personal dose measurements. *Noise & Health*, *18*(81), 98-103.

Wechsler, D. (1997). *Wechsler Adult Intelligence Scale* (3rd edn.), Psychological Corporation, San Antonio, TX.

Wilson, R. H. (1993). Development and use of auditory compact discs in auditory evaluation. *Journal of Rehabilitation Research and Development, 30*(3), 342-351.

Woodcock, R. W. (1987). *Woodcock Reading Mastery Tests – Revised*. Circle Pines, MN: American Guidance Service.

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson III Tests of Cognitive Abilities*. Itasca, IL: Riverside Publishing.

Woodford, P.G. (1996). Evaluating Edwin Gordon's music learning theory from a critical thinking perspective. *Philosophy of Music Education Review, 4*(2), 83-95.

Zendel, B. R. & Alain, C. (2012). Musicians experience less age-related decline in central auditory processing. *Psychology and Aging*, *27*, 410-417.

Zhang, L., Peng, W., Chen J., & Hu, L. (2015). Electrophysiological evidences demonstrating differences in brain functions between nonmusicians and musicians. *Nature: Scientific Reports*, 5:13796.

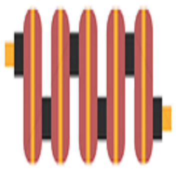
APPENDIX A

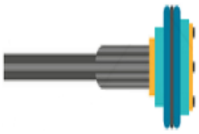
MUSIC TRAINING & LEARNING STUDY POST-TEST

Music Training & Learning Study Post-Test

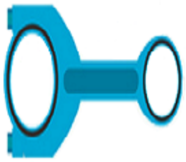
ID:__________________ Version:_____

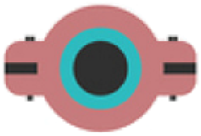
Instructions: *The images below are the same images you learned the names of in the lab. For each picture, please select the name from the word bank below that goes with each picture, and write it on the line below each image.* *There are 20 words in the word-bank, and 10 images; you will NOT use all of the words.*





_______________ _______________ _______________ _______________ ______________





_______________ _______________ _______________ _______________ ______________

mednost	**domfush**	**sentop**	**homtul**	**smedmod**
pedton	**todsun**	**depmost**	**sothnud**	**kentop**
kensom	**stomun**	**foznush**	**doztul**	**nothfud**
stomtul	**tensom**	**sedtoss**	**stoznud**	**smentoss**

APPENDIX B

SPOKEN PASSAGE TRANSCRIPTS

Passage 1

Today's biologic medications are significant to patients with serious or chronic illnesses. A biologic *somnud* is an example of biotechnology. Among other applications, biotechnology may involve the use of a **homtul** to produce a medical treatment. Most biologics come from cells that have been genetically engineered to produce a particular *hoznush*. This process involves introducing **stomun** into a specific type of **homtul**, typically a harmless type of bacteria, yeast, or mammalian cell, which acts as a host cell. The **stomun** tells the cell how to produce the *hoznush*. Once a cell has been engineered, the next step is to create a *fothnul*, which has a complimentary function. This unique *fothnul* is then frozen and stored, and is used as the **doztul** from which all future copies of the cells are made. A biosimilar is a *somnud* that is highly similar to the **doztul**. However, a biosimilar is not considered a generic. Generics are medications that are chemically identical to the original brand-name products. A biosimilar *somnud* is different, however, in that its **stomun** is not identical to that of the original biologic. The **homtul** that is used as the host cell may differ as well. That is, they come from an entirely different *fothnul*. However, each biosimilar must undergo rigorous testing to ensure that the *hoznush* is effectively the same as the original biologic in terms of the bioactivity of the **doztul**. Many patients are hopeful that biosimilars will provide a more affordable treatment option.

Passage 2

FIRST Tech Challenge is a student-centered activity that requires each team to design, build, test, and program an autonomous and driver-operated robot that must perform a series of tasks. Rookie teams are provided with simple, basic instructions for building a functioning robot that can be successful in competition. The robot design is based around a **kensom**, which forms the base of the robot and connects the *pentoss* to the rest of the machine. The **sentop** forms the upper portion of the robot and provides stability to the robot's *temson*. This component also houses the robot's electronics so that it can be driven remotely. If built correctly, the *temson* of the robot can extend to manipulate objects using the **depmost** at its end. The component that makes the robot autonomous is the *pentoss*. The whole machine is powered by the drive motor. For this reason, the **sentop** must be sturdy in order to provide precise movement to the **depmost**, but also lightweight, so as not to overload the motor. Assembly of the **kensom** consists of using socket-head cap screws to connect each *kedton* to form the rectangular base. The *temson* of the robot, when not in use, rests on the front *kedton*. Components used in the construction of the **sentop** include a set of gears and axels that allow the **depmost** to be extended forward or upward. *Kedton* alignment is critical to construction, as the *pentoss* of the robot will be difficult to maneuver if the **kensom** is crooked.