

4-13-2005

Effect of Age on 11- to 18-Year-Olds' Discrimination of Nuances in Instrumental and Speech Phrase Interpretations

Andrew Sioberg
University of South Florida

Follow this and additional works at: <https://scholarcommons.usf.edu/etd>

 Part of the [American Studies Commons](#)

Scholar Commons Citation

Sioberg, Andrew, "Effect of Age on 11- to 18-Year-Olds' Discrimination of Nuances in Instrumental and Speech Phrase Interpretations" (2005). *Graduate Theses and Dissertations*.
<https://scholarcommons.usf.edu/etd/867>

This Dissertation is brought to you for free and open access by the Graduate School at Scholar Commons. It has been accepted for inclusion in Graduate Theses and Dissertations by an authorized administrator of Scholar Commons. For more information, please contact scholarcommons@usf.edu.

Effect of Age on 11- to 18-Year-Olds' Discrimination of
Nuances in Instrumental and Speech Phrase Interpretations

by

Andrew Sioberg

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
School of Music
College of Visual and Performing Arts
University of South Florida

Major Professor: Jack Heller, Ph.D.
C. Victor Fung, Ph.D.
David Williams, Ph.D.
Sheila Woodward, Ph.D.

Date of Approval:
April 13, 2005

Keywords: cognition, interpretation, perception, listening, nuance

© Copyright 2005 , Andrew Sioberg

Table of Contents

List of Tables	iii
List of Figures	iv
Abstract	v
Chapter One - Introduction	1
Background of the Problem	1
Purpose	5
Problem	5
Hypothesis	6
Chapter Two - Review of Literature	7
Perception and Cognition Theories	7
Summation	16
Developmental Theories in Music Perception and Cognition	19
Meter and Tempo Studies	25
Pitch Development	26
Auditory Perception	27
Rhythm	27
Aptitude	28
Interpretation	29
Music and Language	32
Chapter Three - Methods	36
Problem	36
Development of the Items	36
Pilot One	40
Pilot Two	41
Study Population and Sample	45
Variables	46
Measures	46
Data Analysis	46
Chapter Four - Results	47
Assumptions	47
Overall Test Results	49
Speech and Music Part Results	51
Chapter Five - Discussion	58
Limitations	61
Future Study	62

References	64
Appendices	70
Appendix A: Percentage of Students who Answered Correctly Each Item of the Five Speech Categories	71
Appendix B: Percentage of Students who Answered Correctly Each Item of the Five Nonsense Categories	72
Appendix C: Percentage of Students who Answered Correctly Each Item of the Five Musical Except Categories	73
Appendix D: Percentage of Students who Answered Correctly Each Item of the Five Composed Melodies Categories	74
Appendix E: Pilot One: Descriptive Statistics	75
About the Author	End Page

List of Tables

Table 1	Pilot 1 Descriptive Statistics, Cronbach’s alpha, and Pre/Post Pearson Product-Moment Correlation	43
Table 2	Pilot 2: 6 th , 8 th , 10-12 th Grade Descriptive Statistics	44
Table 3	One-Way ANOVA Summary Table of 6 th , 8 th , and 10-12 th Grade	45
Table 4	Age 11, 12, 13, 14, 15, 16, & 17-18 Overall Test Score ANOVA Descriptive Statistics	48
Table 5	7 x 2 x 2 Factorial ANOVA Summary Table Investigating the Relationship Between Age, Gender, and Prior Lesson Experience on Students’ Ability to Interpret Music and Speech Phrase Changes	50
Table 6	Age 11, 12, 13, 14, 15, 16, & 17-18 Speech Items ANOVA Descriptive Statistics	52
Table 7	Age 11, 12, 13, 14, 15, 16, & 17-18 Music Items ANOVA Descriptive Statistics	53
Table 8	7 x 2 x 2 Factorial ANOVA Summary Table Investigating the Relationship Between Age, Gender, and Prior Lesson Experience on Students’ Ability to Interpret Speech Phrase Changes	53
Table 9	7 x 2 x 2 Factorial ANOVA Summary Table Investigating the Relationship Between Age, Gender, and Prior Lesson Experience on Students’ Ability to Interpret Music Phrase Changes	55

List of Figures

Figure 1	Overall Lesson Score Means by Age Group in Percentages	51
Figure 2	Speech Item Score Means for Lesson Experience Arranged by Age Group in Percentages	54
Figure 3	Music Item Score Means for Lesson Experience Arranged by Age Group in Percentages	56

Effect of Age on 11- to 18-Year-Olds' Discrimination of
Nuances in Instrumental and Speech Phrase Interpretations

Andrew Sioberg

ABSTRACT

This dissertation was a continuation of study on a theory of a learning window for the perception of expressive qualities in music and speech. The proposed theory suggested that a practice window must overlap a learning window before it closes around the age of 10. This dissertation sought to determine whether children older than the proposed learning window continued to improve in speech and musical discrimination skill, or leveled off in this ability. It also examined the impact of gender and private lesson experience on discrimination ability.

Instrumental music students ($n = 292$) attending a public magnet school for visual and performing arts in North Carolina between the ages of 11 and 18 participated in the study. Each student was administered a forty-item listening test containing 20 speech items and 20 instrumental music items. Each test item consisted of three short speech or musical phrases. All three phrases in each item were the same written words or notated music, but one phrase was different in interpretation or expression from the other two phrases. Two of the phrases were intended by the performers to be the same in interpretation or expression and one was intended to be different in interpretation or

expression. Subjects were asked to determine which of the three phases in each item was different in interpretation or expression from the other two.

Results of the study suggested that students with prior private lesson experience scored significantly higher than those students that had never taken private lessons. This study seemed to reinforce the proposed learning window for speech and music interpretation in that interpretation ability did appear to level off.

Chapter One

Introduction

Background of the Problem

This study was designed to further test a theory of a learning window for the discrimination of expressive qualities in music and speech. Learning window theory is an established belief that specific time spans occur in a learner's life when certain skills (in this case speech and music syntax) are easiest and most conducive to acquisition (Flavell, 1965). Campbell and Heller (1972) proposed a theory that a practice window must overlap this learning window before it closes around the age of 10. If students are given opportunities to practice music and speech before the closing of this window, children seem to learn the necessary implicit rules for music interpretation and language prosody.

Validation of this theory reinforces the importance of music early in life for all children. If the music perception/learning window operates at a young age, exposure to music within that time period may be critical for students to develop related musical interpretation contingent with life-long musical learning and appreciation. If learning implicit rules to music must occur at an early age for children to develop the ability to perform culturally appropriate music throughout life, it follows that other genres of music deemed of value must also be incorporated much earlier in music education. This may lead to more rigorous study earlier in students' development of jazz, world musics, and other genres deemed valuable by the profession to assure the highest assimilation of implicit rules to those, and other, valued forms of music.

Research in the field of language studies has produced evidence supporting the theory of a learning window being in place for acquisition of language syntax. Nash (1997) stated that the window for acquiring syntax may close as early as five or six years of age, but the window for learning new words never closes. According to Nash, the ability to learn a second language was highest from birth to six years of age, and declined from there. While it is possible for people to learn language after the learning window closes, it is only at great difficulty. Neuroscientists (see Nash) believed the brain's greatest development closed by the age of 10. Nash recommended foreign languages be taught in elementary school or earlier.

As with different spoken languages, different musical 'languages' require learners to know a specific set of implicit rules to properly perform them. Campbell and Heller (1972) suggested this set of implicit rules for music was acquired at the same time syntax and the implicit rules for language were acquired. In 1982 Heller, Campbell, and Gibson demonstrated that this ability seemed to level off by age ten.

According to Heller and Athanasulis (2002), whatever the musical genre or spoken language, the performer drew from specific, implicit stylistic or syntactic rules to deliver the performance in a manner deemed culturally appropriate. In both instances, "how" we perform the musical or spoken phrase was just as important as "what" was said or "what" was played. According to the researchers, these implicit rules were not directly observable, and were therefore not easily defined. However, their presence could be inferred from the resulting musical and speech nuances performers made, given the cornucopia of variation possible in the delivery of each musical or speech phrase. With regard to both music and speech delivery, variation could be observed in the pitch,

timbre, rhythm, dynamics, meter, and timing of a phrase, among others. These possibilities could also occur in any combination and at any point of a musical or spoken phrase. In speech, these subtle variations were called supra-segmentals. In music, the variations were often referred to as nuances. In both speech and music, Heller and Athanasulis (2002) believed the development of this discrimination skill took place at an early age.

In addition to the complexity of the continuous variation in music and speech, Heller and Campbell (1981) argued that traditional research methodology, isolating and observing a single variable, to investigate these implicit rules produced an artificial and inaccurate representation of the music or speech phrase. Their position was that music nuance and speech supra-segmentals were context-dependant and could not be observed independent of the entire, unaltered, audible performance.

The theory Heller and Campbell (1981) supported moved away from the concept of music ability as a special talent, and moved towards the idea that music was an activity all humans were capable of achieving, given access to a musically rich environment by a certain age. In addition, they proposed that perceptual development in music had similar characteristics to perceptual development in spoken language. Music, like spoken language, involved a system by which multiple cues were implemented by the performer to evoke covert responses in the listener. Both were an active process involving the individual interacting with the environment, refining patterns and prominent features. Recognition of invariance to both music and language was significant to development of perception; however, perceptual invariance did not imply stimulus invariance (see Carl

Seashore's "normal illusions," 1938). These normal illusions also put significance on cultural conventions and their importance in perceptual development.

In an effort to investigate this theory, Heller and Athanasulis (2002) developed a 30-item speech test and a 30-item music test and administered it to 1st- 3rd- and 5th graders. The speech test contained short phrases spoken by a child, woman, and man. The music portion contained melodic phrases performed by a clarinet, cello, and piano. In both tests, each item contained three short phrases, with one phrase being different in interpretation from the other two. The task for the listener was to identify which of the three phrases was different in interpretation. The study found that students were better at this task the higher the grade they were in, which does not support the learning window theory. The expectation was that fifth graders' ability (students about the age of 11) would not be higher than third graders (about the age of 9).

In another related study, Heller (2003) investigated the ability of 66 undergraduate non-music majors to take similar tests and compared their results with their musical aptitude scores based on the *Advanced Measures of Music Audiation* (Gordon, 1989). Results showed there was a significant difference on the music items between the subjects who scored high and low in music aptitude. There was also a significant difference on the speech items between high and low music aptitude subjects. Those subjects scoring high in music aptitude scored high on both music and speech tests.

The current study sought to continue to investigate the model of a learning/practice window from birth to about the age ten in order to further examine the viability of the theory. This dissertation examined the ability of students in the sixth to

twelfth grades (ages 11-18) to take a speech and music discrimination test. Continued improvement over time on this test's scores may suggest that the theory needs revision. Sustained scores over the grade levels would reinforce the theory. If there was no apparent growth in student discrimination abilities from grades six through twelve, a reasonable conclusion might be that the skills tested have already been developed by the sixth grade.

Purpose

The purpose of this study, then, was to examine the ability of children older than the proposed critical learning window. It sought to answer the question: Will students in grades 6-12 improve their ability to discriminate changes in speech and music phrase interpretations?

Problem

In order to address the stated purpose, the following questions must be addressed:

1. Can children older than the proposed limits of the learning window demonstrate significantly higher scores in discriminating differences in culturally valid interpretations intended by performers competent in speech and music as age increases?
2. Is there a difference in spoken phrase and instrumental music phrase discrimination ability across age groups?

3. What effect does gender have on the ability of students to discriminate in music and speech phrases?

4. What effect does a student's private music lesson experience have on his/her ability to discriminate changes in music and speech phrases?

Hypothesis

The following null hypotheses were established for statistical testing:

1. There were no significant differences between the mean scores on the spoken discrimination tasks across age.

2. There were no significant differences between the mean scores on the musical discrimination tasks across age.

3. There were no significant differences between male and female test scores on the discrimination test across age groups.

4. There were no significant differences between students that have taken private lessons with those students who have not taken private lessons in their test scores across age groups.

5. There were no interactions.

Chapter Two

Review of Literature

There is abundant literature in the field related to the proposed study. Scholars have been writing about the process of human perception and cognition since the beginning of the 1900s. Many theories and concepts have been proposed, revised, and developed in that time. A number of articles have examined how perception and cognition are developed throughout the lifespan of humans. Researchers have also examined how perception and cognition pertain to specific elements found in music such as pitch perception, rhythm, and aptitude, among others. Work has also been conducted on the relationship between music and speech.

Perception and Cognition Theories

Psychological research and discussion on the subject of music perception have appeared in the research literature for over a century. Throughout the 1900's and into the 21st century, a variety of theories and models have been proposed addressing the process of music perception. Notable theories and models were articulated by Seashore (1919) and his "copy paradigm;" Meyer (1956) and his expectancy theory; Moles (1958) and information theory; Serafine's (1983) "music as thought" and her construction paradigm; Fiske (1996) and his metalanguage; and Gardner's (1983) multiple intelligences. These theories have changed dramatically over the years as continuing research has improved and refined the understanding of the process of music perception. It remains a core issue in music study since an operational definition and understanding of music perception and

cognition is deemed essential to music education, pedagogy, and curriculum design. This chapter will examine significant contributions of the development of perception and cognition theory.

Carl Seashore (1919) developed a theory stating that musical intelligence functions independently of general intelligence. His work was grounded on what has been called “copy paradigm,” (Fiske, 1996) a belief that anything perceived, regardless of the senses involved, is directly related to measurements of an actual, real, physical object. “...everything that is rendered as music or heard as music may be expressed in terms of the concepts of the sound wave” (Seashore, 1938, p. 2). This meant that the emotional and aesthetic qualities found in music listening could be located within the sound wave itself. In response to this logic, Seashore focused on carefully analyzing and describing the sound wave and its acoustic properties believing he was also simultaneously discussing the structure of music perception.

Out of this theory, Seashore developed a sophisticated series of tests, the *Measures of Musical Talents*, to measure an individual’s inherent musical abilities regarding the sense of tone quality, sense of consonance, sense of volume, and sense of rhythm to correspond with the four psychological attributes of sound: pitch, loudness, time, and timbre (Seashore, 1938, p. 2). Success in music was contingent to natural capacities of each ability in each individual, and this could be considered a level of musical aptitude. He believed musical talent was inborn to each individual (Seashore, 1919, p. 6). The theory was misleading, however, when discrepancies occurred in the data collected between the listener’s perception and actual acoustic measurement. They were described away as “normal illusions” because most people perceived the event in

the same manner (Seashore, 1938, p. 17). In addition, copy theory inferred that all listeners have the same internal representation of the music they **are** listening to. If listeners differed in internal representation of the music, one must be wrong because there could be only one correct representation. While Seashore's work carefully described the acoustic signal, it did not accurately represent the perception of the listener.

Meyer (1956) proposed an important theory regarding expectation in music. During the course of a musical work, listeners created on-going expectations about what particular tonal-rhythmic events were likely to occur "next" in the composition. Those expectations were based on events that have already taken place in the music. "Expectation then is a product of the habit responses developed in connection with particular musical styles and of the modes of human perception, cognition, and response – the psychological laws of mental life" (Meyer, 1956, p. 30). Based on past experience, if a stimulus in the present led the listener to expect a future consequent musical event, that stimulus then had meaning. According to Meyer, those stimuli that did not bring expectation had no meaning.

Fiske (1996) believed there was a problem with testing expectancy theory because it did not offer a method for specifying different degrees of expectancy strength either between musical events or for repeated listenings. A partial solution to this problem was an "expectancy profile" offered by Carlsen (1982). A listener's expectancy profile hypothesized more than one pattern progression which may occur "next" in the piece. Fiske (1996) suggested measuring the specific probability of expected events and relative strength of their associated emotional arousals would produce better results.

Developing out of Meyer's expectancy theory, information theory provided a statistical theory towards the development of an understanding of music perception. Fiske (1996) described information theory as the result of a refined sensitivity to the relative probability of particular (musical) events and their future occurrence. Tonal-rhythmic messages sent by the performer to the listener resulted in meaning, but not because meaning existed in the message. Instead, meaning was realized by the listener according to a set of conventions agreed upon by the performer and the listener (Fiske, 1996; Heller and Campbell, 1981). More specifically, Moles (1958) attributed the weakness of studying musical acoustics as its lack of addressing the real problems raised by the creation of sonic structures: "...acoustics studied the rubbing of the bow on the string, when the only thing that interested the musician was the tone this string produced" (Moles 1958, p. 105).

Moles (1958) proposed that information theory defined the complexity of the message being sent from the performer to the listener as the degree of originality. The more original the material in the message, the more complexity it contained. In the theory, original or unexpected events provided information, whereas expected or previously heard events provided no information.

However, there is no art without constraint. To say that music is an art is to say that it obeys rules. Pure chance represents total liberty, and the word *construct* means precisely to revolt against chance. An art is exactly defined by the set of rules it follows. The role of esthetics... is to enumerate these rules and link them with universal laws of perception (Moles, 1958, p. 105).

Davies (1978) suggested listeners understand music, or music has meaning because it confirms or disconfirms certain expectations, which supported Fiske's (1996) discussion of information theory. Expectations and predictions depended on the organization of musical material and the knowledge a person had about the organization of the material. (Davies, 1978).

Neisser (1976) suggested that the listener is constantly developing anticipations of what will come next, based on the information that has already been picked up. It was these anticipations formulated in temporal patterns that governed what would be picked up next, and were then modified by it. The anticipations were not highly specific, and were rarely definite, only guiding the listener in a general direction. Neisser defined cognition as "the activity of knowing: the acquisition, organization, and use of knowledge" (Neisser, 1976, p. 1).

Serafine's theory of "music-as-thought" sought to show that musical comprehension was a result of active cognitive construction and not a passive observation of musical structures. She flatly rejected the idea that meaning was in sound along with any theory suggesting that the analysis of sound would produce musical meaning (Fiske, 1996).

To say that 'the composer inverted the theme' or 'the chord sequence modulated to V' is to say that such themes and chords will be *heard* as inverted or modulated. The results of formal analysis should be consistent with cognitive reality, or they fail as explanations of art. (Serafine, 1983, p. 121)

According to Campbell (1991), Serafine argued that music comprehension required problem-solving strategies involving higher-order cognitive processes not unlike that of chess or mathematics. Serafine also thought music study contained an emphasis on organization in a temporal context and not the physical entities of sound. Field definition, which accounts for music characteristics, temporal organization (event to event processes), and nontemporal operations (formal characteristics of a work) are three cognitive processes she suggested were involved in music comprehension.

Serafine also took a strong opposition to the assumption that musical elements were found in nature. She believed that scales, chords, and music theory in general had nothing to do with music cognition. Instead, they presented a system with which to order and describe music. Fiske observed that Serafine took the position of construction paradigm that musical understanding was:

...a product of pattern generating/reception processes rather than an aural copy of information contained in sound objects. Her theory emphasizes cognitive processes; perceived musical structures as a by-product of musical thinking; the shared cognitive processes of the composer, performer, and listener; and the non-communicative function of music (Fiske, 1996, 51).

Serafine believed that psychological research had been too narrowly focused on the artifacts of analysis than with music. She believed music study must contain an emphasis on organization in a temporal context and not the physical entities of sound (Serafine, 1983).

Taking this constructivist theory even further, Lerdahl and Jackendoff developed a theory that the listener's mind constructed music entirely. Rhythm, beat, and melody were derived from a physical symbol (the sound). The listener's musical experience or "intuition" was the catalyst for the organization of the auditory signal (Lerdahl and Jackendoff, 1983).

Studies by Krumhansl (1983) suggested that single tones within a tonal context were interpreted by the listener according to their function and organization. In addition, chords which were built with related harmonic and melodic information helped to develop the concept of tonality in the listener. Listeners apparently made reference to knowledge of the regularities underlying music within their experience to interpret the incoming sensory information.

Fiske developed the metalanguage theory which addressed the issue of differentiating music sounds as different from other sonic events and also took into account the multifarious musical languages found around the world. According to Fiske (1996), music cognition was a three-stage process. In the first stage, the auditory processing system received any and all incoming acoustic signals and did not yet distinguish music from any other sound. At this point there was only the differentiation between speech and non-speech. The second stage involved a cognitive decision-making activity where pitch-durational patterns were formed and compared to conclude what was being listened to (music, noise, speech). The third stage was the representation of patterns through encoded features for storage, recall, and recognition.

Before a distinction could be made between music and speech, one must have been made between vocal and non-vocal sound through a micro-level temporal

psychoacoustic identification of timbre. Fiske placed this in stage one of his model. He postulated a dual-track processing system in stage two for pattern formation and identification of speech and non-speech sounds using their tonal-rhythmic pattern processing hierarchies. Stage three made stipulations regarding speech-intended vocal sounds and nonspeech-intended vocal sounds and whether they were music or not using the individual's culturally influenced beliefs and experiences. In addition, it made a speech-music stipulation to account for unaccompanied song.

For his theory to work, a number of assumptions were made. Fiske (1996) assumed music perception and cognition was a constructed process and not a copy process. With regard to the musical process, music was a part of all known cultures; was a uniquely human activity; occurred in a wide variety of styles and genres that tended to change over time. Musical activity was about the generation and reception of tonal-rhythmic patterns of sound, and music perception and comprehension required time and effort. Finally, music comprehension was the result of an implicitly known set of cognitive processing rules which could have been either style-specific acquired by experience, or panstylistic inherent in the listening system.

Elliott's (1995) writings supported the work of Fiske and Serafine regarding music cognition and perception, although it was discussed under his definition of listening. He suggested that the act of listening was both constructive and cognitive. A listener's internal organization and experience of music involved several kinds of musical thinking and knowing because the sonic materials being heard were always practice-specific and culture-specific constructions. In addition, music listening was not restricting brain activity to just sonic stimuli, because music was always context-dependent.

Elliott proposed the following:

The procedural essence of music listening consists in such covert, nonverbal acts as constructing coherent musical patterns, chaining musical patterns together, making same-different comparisons among and between patterns, and parsing musical patterns into different types of textures.

(Elliott, 1995, 85)

It is important to note that Elliott's use of the word 'covert' was his attempt at describing implicit or internal actions taken by the listener.

Heller and Campbell (1981) rejected the idea that music contained meaning.

Instead, they assumed a construction paradigm which argued that listeners construct their own perception (assign their own meaning) of sound patterns in the brain.

Heller and Campbell's theory could be organized into a number of features. First, the performer produced a sound which provided the listener with raw acoustic cues. Next, those cues were analyzed for micro-level patterns which were linked together into longer patterns that eventually constituted musical elements. In the third, pattern identification **was** guided by an active, context-dependant, and implicit set of rules by which both the performer and the listener abide. The rules both parties follow were formulated through a social/cultural contract. Like the comprehension of different spoken languages, different musical 'languages' required the participants to know different sets of rules to successfully interpret them. This musical perceptual ability must be acquired early in an individual's development. Heller and Campbell suggested that for discriminating "interpretation" in music nuance, this acquisition occurred by the 10th year of life. Their theory stated that language perception for prosody was also developed by the 10th year.

“Critical periods cannot be recovered; lack of appropriate musical experiences at the time during which the brain is most prepared to acquire music social/cultural contracts results in a musically handicapped adult” (Fiske, 1990, p. 7).

Heller and Campbell’s theory sought to avoid the concept of musical listening ability as a special talent, and instead presented a case to see music listening skill as an activity all humans are capable of achieving, provided they are exposed to a musically rich environment by a certain critical time period. If a listener missed this opportunity within the critical time period, music listening skills and implicit rules to musical interpretation could “...only be developed with great difficulty (or not at all) in a later period” (Campbell and Heller, 1981).

The theory proposed that perceptual development in music had similar characteristics to perceptual development in language. Both were an active process involving the individual interacting with his or her environment, refining patterns and prominent features. Recognition of invariance in both music and language was significant to perception. However, perceptual invariance did not imply stimulus invariance (Seashore’s “normal illusions,” 1937). Heller and Campbell also put significance on cultural conventions and their importance in perceptual development.

Summation

Seashore’s work focused on the analysis of the sound wave and the components which comprised it. No matter how careful the analysis was or how specific the information was describing the sound wave, this method of study did not adequately address the emotional or aesthetic qualities each person derives from a musical experience. An individual’s response to music is not measured by components of the

sound wave. Seashore's work helped to contribute to the study of music perception and cognition, but his theory was not directly applicable to this study.

Meyer's proposal that the listener created continuously evolving expectations as music was performed comes closer to a functional address of perception and cognition. The theory falls short, however, when Meyer suggested that stimuli which do not bring expectation (those musical events which have not been experienced before) have no meaning. Artful music created over the centuries has continuously challenged conventions of the time. Composers have skillfully and intentionally crafted their music in such a way as to delay or even defy expectation specifically to arouse meaning in the listener.

At the same time, this position does not imply that music with little or no adherence to the listener's conventions would somehow have more meaning. Moles developed a similar position by suggesting that original or unexpected events in music provided information whereas previously heard material provided no new information. If this were the case, there would be no need for previously experienced musical events in a composition.

An important component of meaningful music, then, is a careful balance of new material intertwined with the known, traditional conventions of the listener. This delicate balance of interest (new material) versus clarity (known material) continuously evolves as the piece progresses, for as new material is introduced, it loses its novelty with time. Neisser (1976) and Serafine (1982) both allude to this point. In music, original events and those events conforming to convention should be perceived as equally important in the listener's perception and cognition process. Music cannot be understood without both.

Serafine (1983), and Lerdahl and Jackendoff (1983) provided another important component to music perception and cognition by proposing that meaning in music was constructed by each listener according to his/her individual listening experiences. The listeners' construct, however, must be guided by some supervisory, decision-making process. Conventions, as discussed by a number of scholars (Moles, 1958; Fiske, 1996; Heller and Campbell, 1981) were developed over the life experience of the listener and performer. Heller and Campbell's (1981) proposal of culture was the most logical determinant for assembling this convention. Fiske (1996) enhanced this influence of culture on music perception and cognition by also allowing for processing rules to be inherent in the human auditory system. Hints of the auditory system having an impact on the listening process can be seen as far back as Seashore's (1938) work and his identification of inherent musical abilities.

Heller and Campbell (1981) took the music learning process one step further by seeking to match its cognition processes to that of the speech learning process. Their theory proposed that musical interpretative ability was developed about the same time speech interpretative ability was obtained. Furthermore, both skills were acquired the same way; in an active process employing a recognition of invariance. The current study was developed to further examine the validity of this proposed theory.

Developmental Theories in Music Perception and Cognition

Scholars have long been examining the developmental processes of perception and cognition. Understanding when these skills are acquired is critically important to developing sound, productive educational pedagogy.

The cognitive developmental psychology of Piaget is well respected and has had significant impact on music education (Pfleiderer, 1967; Gardner, 1983; Campbell, 1991; Zimmerman, 1992 and 1986). In his theory, schemata continuously unfold in systematic stages as a result of environmental interaction. Development is continuous and builds on that which has already been experienced. Age levels at which the changes occur vary depending on each individual's cultural, physical, and social environments.

Piaget carefully described the progressive development of the learner as a series of stages. In the first eighteen months of life, intellectual processes are initially developed through predominantly sensorimotor behavior. In the second, or preoperational representation stage, a child undergoes a preconceptual phase and an intuitive phase from about eighteen months to seven or eight years of age. Here the child reasons transductively, still relying heavily on perception for thought. Failure of transductive reasoning causes intuitive regulations and the formation of categories. Continuous and systematic variations in the child's focus of attention provide more impressions of reality in a move towards operational thought. By the age of seven, concepts acquired to this point are organized into coherent systems. Thought is no longer reliant on perception. In the formal operations stage from 12 to 15 the child develops reasoning which culminates in reversibility, the formation of a thought which can then be "unthought" (Flavell, 1965).

A number of researchers have applied Piaget's developmental theory to music and arts. Wolf and Gardner (1980) proposed four stages of artistic development corresponding to Piaget's stages of cognitive development: (1) child as direct communicator, (2) child as symbol user, (3) youth as craftsman (covers 5 or 7 to 11 or 13), and (4) youth as critic and full participant in the artistic process.

Pflederer (1967) proposed that processes involved in musical intelligence stem from perceiving, comprehending, and organizing the structure of music. Music learning begins with perception and the sound structure, followed by musical concepts which permit one to think about what has been heard. "The essence of musical intelligence is found in the framework of rhythmic, melodic, harmonic, and formal relationships which have developed through a progressive organization of musical experiences" (Pflederer, 1967, p. 221).

Pflederer cited five conservation laws in the development of musical concepts. Identity applied to the maintaining of thematic materials over time. Metrical grouping represented an active organization of tonal stimuli around accents. Augmentation and diminution involved the presentation of a musical subject in doubled and halved values which requires a relational framework based on Piaget's reversible operation. The invariance of tonal organization was a conservation law employing transposition. Lastly, inversion of either melody or harmony used the reversible operation, a principle of conservation.

Zimmerman (1982) suggested that processes involved in musical intelligence stem from perceiving, comprehending, and organizing the structure of music. Assessing comprehension was accomplished by observing the cognitive behavior of the listener.

Her discussion was very closely related to the work of Piaget. According to Zimmerman, learning processes begin with immediate and limited perception due to the limited framework for music in the child. Experience moves immediate perception to mediated perception, or a change from direct observation and stimulation to perception implied by or derived through something else. The initial percept moves to the emerging concept. Expansion of experience paralleled expansion of cognitive framework. Concepts became interlocking and relational, allowing for advanced cognition such as transposition, modulation, inversion, augmentation, and diminution.

Sloboda's (1985) views also seem to parallel Piaget's work. He proposed that human music skill was built out of a base of innate abilities and tendencies, and that the physical world and a person's experience in it develop this skill. He divided his discussion of musical learning and development into two main categories: enculturation and training.

With regard to enculturation, there exists a set of inherent capacities present in all humans at birth. There is also a set of shared experiences culture provides children as they grow. Finally, there is the impact of the rapidly changing general cognitive system as culture-supported skills are learned. These factors produced a relatively similar sequence of development for most children within a culture, and a set of similar ages in which growth takes place. In addition, enculturation is achieved largely without a conscious effort to learn.

Training, on the other hand, is a self-conscious effort to become more accomplished at a specific skill using explicit instruction. Individuals focus on specific experiences not shared by all members of a culture. While enculturation is the dominant

process of learning up to the age of 10, training takes an increasing importance following the age of 10. (Sloboda, 1985).

Serafine (1988) proposed a theory of music cognition suggesting that musical thought is a result of the posing of artworks embodying finite and organized sets of temporal events described in sound. Temporal organization (succession and simultaneity), and nontemporal operations (closure, transformation, abstraction, and hierarchic levels) were defined as two categories of cognitive processes indicative of how music is understood. Sound is processed according to a hierarchy of structural elements. She tested her theory using a series of temporal and nontemporal studies incorporating musical tasks given to 15 subjects each at ages 5, 6, 8, 10, 11, and adult.

Results indicated that ten- and eleven-year-olds are in possession of most temporal and nontemporal processes. Interestingly, eleven-year-olds showed a slight decrease in the pitch discrimination task, idiomatic construction task, and the motivic synthesis task. Serafine was expecting continued increases in performance throughout. Eight-year-olds could perceive hierarchic levels in simple melodies, easily identify simultaneous combinations of timbres, and discriminate random melodies on a similar level to older students. Most students had difficulty with the rest of the tests in this age group. Because five- and six-year-olds demonstrated almost none of the processes, Serafine surmised that the years from eight to ten represent a period of rapid growth in music cognition. It was also noted that subjects with instrumental training fared no better than those subjects with no training. Her studies suggested that nontemporal, formal, and abstract processes appeared to develop earlier than temporal processes.

Gardner's (1983) work suggested there exists several relatively autonomous human intellectual competences he referred to as human intelligences, one of which he identified as musical intelligence. He proposed that musical intelligence uses three systems of knowing to address symbols (making, feeling, and perceiving). Symbols in music can be either denotational, expressive, or both.

According to Gardner, from birth to 5-7 years of age the child develops from responding instinctively to relating actively to the environment through making, feeling, and perceiving representations. Instinctive reflexes evolve into symbol use and the three systems differentiate. Infants sing and babble, produce undulating patterns, emit individual sounds, and imitate prosodic patterns and tones sung by others. Over time, the child's spontaneous vocalizations develop into echoing and reproducing song fragments. At the end of this stage, children can be heard inventing songs and singing songs from their culture. From age five to ten, elements of different symbol systems become linked to specific activities within domains. Children focus on learning rules, and have an interest in developing technique. School music expands their song repertoire and knowledge base. After ten, children demonstrate continued skill development and integration of the three systems (making, feeling, and perceiving). Critical acumen and critical reflection also develop, providing new insight towards style and music interpretation along with evaluation skills. Gardner's work contributes much to the theoretical cognitive processes involved in human development and music. It serves as a guide for developing curricula and musical activities for the developing student. (Campbell, 1991).

It seems there is consensus in the academic music community that cognitive development is continuous, compounding, and moves through stages. Many of the music scholars discussed have been clearly influenced by the work of Piaget. The theories offered here suggest how and when music concepts are presented is very important to the success of students' acquisition and retention of those concepts. In a number of these theories, (Flavell, 1965; Sloboda, 1985; Serafine, 1988; and Gardner, 1983) scholars propose a shift in cognitive development does occur around the age of 10-11. Flavell (1965) in his discussion of Piaget's work proposed the developing mind transitions from perception-reliant cognition to reasoning and reversibility. Sloboda (1985) saw a transition around this time period from learning unconsciously through enculturation to learning through formal training. Serafine's (1988) research suggested most of the temporal and nontemporal processes were developed by this same age. Gardner (1983) suggested advanced integration of the three systems (making, feeling, and perceiving) along with critical acumen and critical reflection developed at this same stage. In the proposed theory to be tested in this study, language perception for prosody and music perception for nuance have been developed by this same time period. Following this time period it is possible that interpretative skill can be acquired, but only at great difficulty and perhaps never to the level of a person who acquired it at an earlier, more appropriate age.

A number of studies have been conducted over the years examining different specific music skills and their development through different age groups of subjects. The following is a review of selected literature organized by music skill. While these studies are related to the present research in that they are trying to develop a better understanding

of music cognition and perception, the approach of isolating specific components of music for observation and study may not be the best method to observe these developments. The current study seeks to leave the entire music and speech phrases intact as they were originally recorded and observe children's ability to interpret changes in them. It is thought that keeping the musical and speech excerpts in their original, complex forms makes the task observed more realistic to musical listening and more generalizable to everyday musical activity.

Meter and tempo studies

To investigate the development of children's ability to conceptualize and identify meter in music, Jones (1976) had 66 children aged 6-12 years perform 11 musical tasks in increasingly difficult order. His results supported Piaget's theory that children's concept of time develops through three stages. The first stage was characterized by basic perceptual discriminations. The second stage, beginning after seven years of age, was characterized by a conservation of velocity, immediate seriation, and some difficulty with double seriation. Jones defined seriation following the work of Piaget as the concrete operation that involves ordering stimuli along a quantitative dimension (such as length). Double seriation was an understanding of symbols or objects that involve more than one meaning. By about age 9, there was immediate double seriation, a grasp of simultaneity and succession, operation inclusion, and measurement of physical time. The meter concept in Jones' study developed between the ages of 9 and 10 with many children between 9 and 12 having some difficulty understanding meter concept and meter.

As a follow up, Perney (1976) administered Tasks I-V from Jones' study to second and third graders. Perney found there was no significant difference between

students who played instruments and those who did not, and females fared better than males in performing tasks. In addition, there was a relationship between performance on musical tasks and verbal ability for both sexes.

Miller and Eargle (1990) examined the effects of age and musical background independently on simple tempo discrimination. The three age groups represented were middle childhood (7-11), early adolescence (12-15), and adulthood (18-33). They found that subjects' perception of changes in tempo during a piece was more sensitive to general developmental differences than musical training. However, acknowledging the maintenance of a standard tempo seemed more affected by musical training than age.

Pitch development

Davidson and Scripp (1985) asked children aged five to seven in a longitudinal study to write down a song they knew "so that someone else could sing it" in an effort to capture evidence of those musical features they thought were most important in remembering how a tune goes. In those pictorial representations, they found development trends in children's use of symbols, level of sophistications of their notations, and selective attention to musical features depending on the context. As the representations became more sophisticated, pitch emerges as the most robust component of cognitive musical development. The researchers speculated that as pitch development occurred with age, it increasingly became independent of language, kinaesthetic, or number skills. They recommended more focus on the success of musical understanding children have of music through creative symbolic representation.

Auditory perception

An investigation of a learning sequence for music listening skills associated with the detection of alterations in timbre, rhythm, melodic pitch patterns, and harmony was conducted by Hufstader (1977). The project examined almost 600 first, third, fifth, and seventh grade students from four school districts and their ability to take a test of aural perception skills designed by the researcher. Timbre mean scores reached an arbitrary criterion level for all four districts initially by first grade. Rhythm was second at fifth grade for all four districts. Melodic pitch was achieved by two districts at fifth grade, and the other two at seventh grade. Harmony mean scores only reached the criterion level in two of the districts. This suggested an order to the development of specific listening skills used in auditory perception.

Petzold (1963) investigated the development of auditory perception, phrase learning, and melodic reproduction using varying harmonies, timbres, and rhythmic ability. He found that age was a major factor in the development of auditory perception. For most tasks, a plateau was reached by the age of eight, and the most development occurs between the ages of six and seven

Rhythm

Cox (1977) examined almost 800 students in grades 1-6 from 48 different classrooms in 8 schools to develop a descriptive analysis of their responses to beat, meter, and rhythm pattern. The results of her beat, meter, and rhythm pattern tests indicated that music concepts develop gradually, rhythmic ability increases continuously with age, and the concept of meter does not mature before the age of 9 to 9½.

Atterbury's study (1983) involved ten rhythm tests divided into three categories given to 20 seven-year-olds and 20 eight-year-olds, half were normal achieving readers, and half were learning disabled readers. Results indicated that learning-disabled readers perceived simple same and different rhythm patterns similarly to normal readers, but were less successful at difficult rhythm patterns. Learning-disabled readers also could not reproduce rhythm patterns as well as normal readers. In addition, tapped and spoken rhythms were significantly different than melodic or tapped for all subjects. Atterbury suggested including rhythm syllables when teaching rhythmic patterns to seven- and eight-year olds.

Palmer (1990) engaged fourth-grade students from three elementary schools to examine the effectiveness of the methods of Mary Helen Richards and Edwin Gordon compared to a control group. The study showed that both Richard's and Gordon's methods produced significant gains in rhythm reading over the control group, with Gordon's method showing slightly better results than Richards.

Shehan (1987) examined the effects of aural and visual approaches to rhythm reading and short-term retention with 2nd and 6th graders. She found that the simultaneous use of auditory and visual channels facilitated learning of rhythms in both groups. More importantly, however, the older group was able to learn rhythms twice as fast as the younger group regardless of the presentation method. In this study, age improved the musical skill of rhythm reading and short term retention.

Aptitude

Using Gordon's Primary Measures of Music Audiation (PMMA), Flohr (1981) examined the influence of short-term music instruction on five-year-old children's

developmental music aptitude. Twenty-nine children were placed in three groups consisting of improvisation instruction; singing, playing percussion instruments, and movement; and no music instruction. Results suggested students receiving either form of musical instruction scored significantly higher on the PMMA than the control group.

De Yarman (1975) conducted a study administering the *Musical Aptitude Profile* for four years to fourth graders with varying training and found that the type or amount of formal musical training had little effect upon children's musical aptitude prior to fourth grade. In addition, musical aptitude stabilized before age five or six and was resistant to further training. Schleuter and DeYarman (1977) did a follow up study which further supported the De Yarman's initial findings.

Interpretation

A few studies addressed specifically the issue of music interpretation. Johnson (2000) analyzed 24 piano performances of a portion of Beethoven's Symphony Number 5 in c minor seeking to determine what rhythmic and dynamic variations were used. Eight piano performing experts were asked to perform three times. The first time was played as written, the second time they added their interpretation to the performance to make it as musical as possible, and the third time they exaggerated their interpretation from the second effort to the point the music was unmusical. Results indicated that while dynamic changes were rather small for each take, changes in timing were much more dramatic. This suggested that in piano performance, the performer contributed more to interpretation using timing variations, and relied more on the score for dynamic changes.

Gibson (1986) examined the ability of 4- to 6-year-old children to perceive changes in culturally identifiable interpretative differences of speech and musical

performance. Visual stimuli were used first to determine if subjects could group categories of line drawings. If subjects were successful at this first task, they proceeded to group categories of spoken, sung, and played interpretations. Gibson found an increase in mean score performance for each successive age group. Subjects seemed more adept at categorizing visual stimuli than auditory information.

Rodriguez and Webster (1997) conducted research to determine the nature of children's verbal responses to repeated hearings of a brief musical excerpt. Thirty-three kindergarden through fifth grade children gave responses to questions which were recorded and submitted to three judges who categorized the responses into music listening or other developmental studies. They found similarities between responses and the recognized developmental theory of Gardner (1983).

Denardo and Kantorski (1995) conducted a study that investigated second and fifth graders' musical cognition when listening to four-phrase songs. The children were asked to determine if the second, third, and fourth phrases were the same as, similar to, or different from the first phrase. There was no significant difference between the two groups in ability, accuracy declined the further away in time the phrase in question was to the initial reference phrase, and subjects were more accurate in identifying phrase types that were different followed by similar (one parameter change), similar (two parameter changes), and same. This may support the learning window theory.

Holahan, Saunders, and Goldberg (2000) examined college musicians, non musicians and 1st grade children to investigate development of tonal audiation skills and test the validity of a previous study conducted by Holahan and Saunders in 1993. Their work found the three groups differed from each other in regard to accuracy and speed of

response. College musicians demonstrated systematically higher scores than both nonmusicians and first graders. First graders tested only slightly lower than nonmusicians.

McBeth (1992) described music interpretation as the most difficult of all pedagogical efforts because it **does** not fall into the category of technique. Instead, it could only be achieved through a thorough understanding of the intent of the composer. “Correct interpretation is the re-creation of the composer’s intent” (p. 17). This reinforced from a conductor’s perspective the learning window theory about the cultural relationship between performer and listener.

Shaffer and Todd (1994) examined the perception of rhythm and timing by analyzing the nuances of rhythmic patterns in repeated piano performances using actual performances of pieces. In this study, it was found that pianists slow down the tempo at places of stability in a very precise manner. At the cognitive level, this implied that the motor system was accessing a stable timekeeper and had a definite representation of the relevant timing parameters. At the musical level, it meant that tempo and rubato were not left to chance, but were the outcome of coded decisions, or implicit rules, of interpretation.

Nettl’s work (1994) supported the theory that music is culture bound. He suggested music expresses a subtext whose message is determined by extraneous issues such as culture, class, gender, and personality. While he was unable to identify explicitly musical thought as different from other kinds of thinking, he did suggest the way musicians think about their music depends largely on the way they perceive their world at large. The way a society thought about the concept of music, shaped the way its musicians made music.

Music and Language

Levman (1992) cited three major positions regarding the origins of music and language. The first was that they both are separate and different faculties. This view supported the notion that language's roots were primarily gestural in origin, the result of a primitive representation system early humans developed to navigate the environment. It began with humans mimicking each other along with other sounds in their surroundings. Initial word-sounds were emotive and expressive declarations of species identification, warning calls, cries for help, etc.

The second position was that music developed out of language. Song was the result of emotional speech intensified and systematized. The idea was put forth by Spencer in 1857. Most musicologists of the nineteenth and twentieth centuries subscribed to one of these two positions. John Blacking (1973) believed that music was a species-specific biological human impulse, separate from language, but also inseparable from the social context in which it develops. Nettl's position proposed that both were inseparable as a form of communication and then grew apart (1983).

In the third position, Levman (1992) suggested both music and language evolved out of a common 'proto-faculty' which was primarily musical in nature, out of impulse of the human to survive. The frequency component contained the most sound perception. Frequency contained not only pitch, but timbral quality and vocal information.

The concept of double articulation (Levman, 1992) separated music from language which until then was closer to music than speech. Double articulation was the sonic division of a language into individual phonemes. By themselves, phonemes had no

intrinsic meaning, but in combination they formed morphemes. It was at this point where language ceased to be isomorphic and became symbolic.

London (1996) proposed there was a cultural belief that music was a kind of language. According to London, music is treated as a linguistic phenomenon. There is a tendency to use the language framework to address all meaningful communicative behavior.

Since the 'music is language' framework is so well established in our musical training and the language we use to describe music, it becomes wholly transparent - for many listeners music becomes a subclass of linguistic phenomena. (p. 51)

He contended that, if there was agreement that there was some sort of communication between composer and listener, in such cases the listener relied on his linguistic competence to interpret it.

Fiske (1990) perceived the argument to have two points. The first was whether or not music was processed cognitively in a similar fashion to language. The second was whether or not music had content and whether that content was specifically communicated as in language.

Swain (1996) posed a paradox in musical semantics and established a case for similarity between music and language. According to Swain, music seems full of meaning to listeners from beginning to advanced, yet there is no consensus to specific meaning, a component essential to natural language. Music and language both have varying degrees of semantic potential, but both are limited by the context in which they occur. The semantic range of a sentence is comparatively small because its component

parts define its context very well. Yet, when not in sentence context, individual words have much more range. While the semantic range for an instrumental music passage is even greater still, Swain argues it is only in degree and not in kind. So while listeners often disagree on the symbolic import of a piece of music for lack of context, the same can happen in language. Fiske agrees with Swain that the issue is based on semantics, and adds that one definitive characteristic in language not found in music is that language is denotative.

Heller (2000) continued refining his work in a paper designed to further examine the link between speech's "supra-segmentals", or the inflection and nuances superimposed on words, with the subtle components of expression or interpretation in music (stress, attack, vibrato, crescendo and others). Defining interpretation as "the expressive characteristics of performance," Heller suggested similarities existed between prosody in speech and music interpretation (p. 77).

To access this covert music listening process, Heller rejected the idea of breaking the acoustic medium down into traditional measurements (frequency, amplitude, duration) because reduction of the full sound removed the critically needed context and culture the listener relied on for understanding and comprehension. His work regarding music interpretation mirrored that of Rommeveit (1974) who proposed language was an evocative process involving the shared cultural experience of both the performer and the listener.

The test for this study was based on functional and not acoustical similarity since the later has not proven a good indicator of listener response to interpretative gesture in music. Specific acoustic variables were not needed to predict listener response to music.

The issue stressed was whether the listener could hear the musical interpretation of a phrase that was repeated (but not identical), yet was intended to be the same “interpretation”. The main determinant of the listener’s selection was whether the phrases had enough characteristics in the repetition to be considered “different” or “not different” in interpretation. If the performance was successful, the listener should be able to categorize the interpretative intent of the performer.

Chapter Three

Method

Problem

This study was designed to further examine relationships between the perception of performed music and spoken language as reflected in a model of a learning/practice window that begins to close between the ages of 6 and 10 for the perception of music interpretation (subtle changes in the expression qualities of performance).

Development of the Items

Questions for this study were developed from previous work conducted by Heller and Athanasulis (2002). In their process, three professional performers (a pianist, clarinetist, and cellist) recorded six short melodies specifically notated for interpretation. In each melody, three different interpretations were recorded. To accomplish this, each performer imitated the performed phrase as interpreted by the previous performer. Initial interpretation of the phrase was induced by the notation. For example, the cellist might start the sequence by playing melody one (with interpretation "A"), the pianist and then the clarinetist would follow, intentionally trying to imitate as closely as possible the initial rendition of the performer who preceded them. This sequence was repeated until all three performers agreed that at least three of these performances were good interpretative matches. The same procedure was followed for melody one with interpretations B, and C (each different from the other). This process was continued until

three different melodies, with three interpretations each, were recorded to the satisfaction of the three performers. This produced a total of 162 phrases (6 melodies x 3 interpretations x 3 repetitions x 3 instruments). All of these phrases were digitized and stored in separate computer files. The recordings were sampled at 18,500 bytes/second.

The speech stimulus recordings were developed in a similar manner. The first stage required that three speakers (a male, a female, and a child) recorded six specifically written short phrases (three were nonsense syllable phrases, the other three were meaningful short phrases) each with three different interpretations. Each phrase was about four seconds in length. Each speaker imitated the spoken phrases as interpreted by the previous speaker. Initial interpretation of the phrase was induced by the written words using italics for stress points. For example, the male speaker might start the sequence by speaking phrase one (with interpretation 'A'); the female speaker would follow, trying to imitate as closely as possible the rendition of the male speaker. This was then followed by the child's rendition trying to imitate the female utterance. The order of sequence was randomly changed for each written interpretation. This process was repeated until all agreed that at least three of these utterances were good interpretive matches. The same procedure was followed for phrase one with interpretations B, and C (each different from the other). This process was continued until all six phrases, with three interpretations each, were recorded to the satisfaction of the speakers and researchers. As with the music phrases, the speech procedure produced a total of 162 short speech phrases (6 phrases x 3 interpretations x 3 repetitions x 3 speakers). These phrases were also digitized and stored in separate computer files. The recordings were sampled at 18,500 bytes/second.

The next portion of the test-making process consisted of organizing the music and speech recordings into usable items to be presented to the subjects. Three renditions of a melody or utterance, two with the same intended interpretation, one with a different intended interpretation, were transferred from their separate computer files into a new file. Each of the three phrases in the file was a different playing, but two were intended to be within one interpretative category (e.g., Interpretation A, Interpretation B, Interpretation A'). Each phrase was separated by one second of silence. Following the three initial phrases, 1.5 seconds of silence and the word 'again' followed by another 1.5 seconds of silence was added. The three phrases were then repeated (A...B...A.... 'again' A...B...A). Each collection of three repeated phrases and silence was designated one test item. The length of each test item, three phrases repeated, was about 30 seconds. The task for the listener was to determine which interpretation did not belong with the other two and to indicate their choice on a bubble sheet.

In Heller and Athanasulis' (2002) study, a similar test was created that included thirty items constructed of 30 music items, and 30 speech items. No consideration was given to determine which combinations of phrases would be most effective in testing listeners. For this dissertation, two pilot studies were designed specifically to examine the quality of a wide range of possible test items to help determine which items would be most effective in a final testing format.

In each of the two speech categories (speech phrases and nonsense phrases), items were created five different ways. Items were created using only the child's voice for three interpretations; only the woman's voice for three interpretations; only the man's voice for three interpretations; a combination of child first, woman second, and man last; and a

combination of woman first, child second, and man last. In each of these five subcategories, 18 items were constructed representing a wide variety of the possible combinations. Ninety items were developed for each of the two speech categories. The speech portion contained a total of 180 possible test items.

In each of the two music categories (musical excerpt phrases, and composed music phrases), items were also created five different ways. Items were created using only the piano for three interpretations; only the clarinet for three interpretations; only the cello for three interpretations; a combination of piano first, cello second, and clarinet last; and a combination of cello first, clarinet second, and piano last. In each of these methods, 18 items were constructed representing a wide variety of the possible combinations. Ninety items were developed for each of the two music categories. The music portion contained a total of 180 possible items to choose from.

The 18 individual items generated in each subcategory were then linked together into one listening presentation. Each item (with the three repeated phrases and silence) was separated from the next by 1.5 seconds of silence; a vocalization of the number of the item in the order it was presented; and 1.5 more seconds of silence. Linked together, the 18 items varied in total length from 6 to 11 minutes, depending on the length of the individual phrases which comprised each item. A total of 20 listening assessments (ten in speech and ten in music) each containing 18 items was created out of the subcategories.

The specific listener subject task was to decide which phrase in the three-phrase item was different in interpretation from the other two. In each item all three phrases were different performances (utterances) of the same phrase. That is, phrase one was followed by two more performances of that phrase, but one of the three phrases was

different in interpretation of the same melody or written sentence. The phrase that was the 'same' interpretation was never a re-recording of that interpretation. It was always another performance, but one that was intended by the performer or speaker to be the same interpretation. This makes the task more difficult, yet more generalizable to a realistic musical or speech situation. In a normal musical performance or in normal speaking the performer may intend to repeat specific interpretations of a phrase, but each repetition is not identical. There is variability among phrases that are intended to be the same in interpretation. Musical or language decisions are made by the listener to interpretative category rather than to interpretive equivalence. So the listener's task in this study was to determine which of the three phrases was intended to be different from the other two in interpretation.

Pilot One

The pilot assessments were administered to two sixth grade string classes in an urban, public middle school in the southeast United States. Class 'A' contained between 15 and 19 students depending on attendance. Class 'B' contained between 20 and 25 students. Permission was granted by the string teacher, who incorporated the assessments into her daily class routine as a listening task. Only one assessment was given in any class period. Both classes were given instructions and examples to practice on together as a group prior to their first effort to insure they understood the task. To keep students interested throughout the duration of the assessments, the classes artificially competed against each other for the title of 'Best Listeners', and a period of recreation with snacks was provided to the class with the highest average. Students marked their answers on a scantron and turned in their work without their names on the papers to assure anonymity.

Assessments were taken over a period of six weeks depending on the school, class, and researcher's schedules.

Analysis was conducted to determine the percentage of students who successfully answered each item. Lists were then compiled for each item in each category (i.e. child speech, cello composed melody) with their percentage of success. Appendixes A, B, C, and D show the complete results for speech, nonsense, melodic except, and composed melody categories respectively. Those two items closest to 50% success in each category were selected for use in the study. In situations where more than two items were available, the two chosen were randomly selected from the available options. In categories where two items were not at the 50% mark, the two closest (either high or low) were selected. This effort produced 2 items of relatively similar difficulty from each of the 20 categories for a total of 40 items to be placed in the final study assessment tool. Item selections to be incorporated into the final assessment tool for each subcategory are also presented at the bottom of Appendixes A, B, C, and D. Appendix E lists each subcategory's descriptive statistics including N, mean, percentage correct, median, and standard deviation.

Pilot Two

The second pilot study examined the reliability of the final study assessment tool. For this pilot, the 40 selected items were placed into two randomly generated orders. Both forty-item assessments (A and B) were then administered to two new groups of 6th graders in a different school with similar socio-economic conditions whose music program contained students with similar musical experience. In addition, assessment A was also given to one eighth-grade class and one 10-12th grade class. As with the

previous study, permission was granted by the two music teachers, who incorporated the assessments into their daily class routine. All classes were given instructions and examples to practice on together as a group prior to their first effort to ensure they understood the task. Assessment A was conducted on a different day than assessment B. The order was altered to prevent students from recalling the first test by simple repetition.

Students marked their answers on a scantron sheet in a similar manner of the previous pilot, but in this study, the students were assigned a scantron sheet with a number written on it. This allowed the researcher to identify each student's work on both assessments (A and B) while still maintaining anonymity.

Descriptive statistics, *Cronbach's alpha*, and *pre/post Pearson Product-Moment correlation* of the two 6th grade groups taking both assessments A and B are displayed in Table 1. The *Cronbach's alpha* score of both groups suggest a good internal reliability, reinforcing the item selection process of the first pilot. Scores of .83, and .89 from the two groups examined are well within the established range of acceptability (Nunnally, 1978). *Pre/post Pearson Product-Moment correlation coefficients* of .42, and .58 suggest a small to moderate relationship between test A and test B. While the *Pearson Product-Moment correlation coefficients* are lower than expected, small sample sizes may be allowing for wide confidence intervals that may account for this small relationship. The high *Cronbach's alpha* scores suggest good internal reliability allowing the study to proceed.

Table 1

Pilot 1 Descriptive Statistics, Cronbach's alpha, and Pre/Post Pearson Product-Moment Correlation

Group 1					
Test	<i>N</i>	Mean	<i>SD</i>	<i>Cronbach's alpha</i>	<i>r</i>
				Assessment A & B	Assessment A & B
A	23	.59	.16	.83	.42
B	23	.58	.14		
Group 2					
Test	<i>N</i>	Mean	<i>SD</i>	<i>Cronbach's alpha</i>	<i>r</i>
				Assessment A & B	Assessment A & B
A	18	.54	.17	.89	.58
B	18	.51	.15		

Further examination of an 11-12-year-old group, a 13-14-year-old group, and a 16-18-year old group was investigated using a one-way ANOVA, between groups design to determine if the higher grade improved test results. Descriptive statistics are displayed in Table 2. All groups showed within acceptable skewness and kurtosis with one exception. The 10-12th grade group demonstrated a notably leptokurtic distribution. The 10-12th grade group also had one extreme outlier.

Table 2

Pilot 2: 6th, 8th, 10-12th Grade Descriptive Statistics

	11-12-year-olds	13-14-year-olds	16-18-year-olds
<i>N</i>	19	17	29
Mean in %	55.28	65.45	73.20
S.D. in %	17.13	13.80	15.38
Skewness	-.61	-.42	-.73
Kurtosis	-.53	-.15	1.44

Sample means of 55.28%, 65.45%, and 73.2% for the 6th graders, 8th graders, and 10-12th graders respectively, show enough variation to be of interest. To provide a more standardized measure of effect size, *Cohen's f* was calculated to be .486. This can be interpreted to mean that the group means typically deviate from the grand mean by about .4 standard deviations. An effect size, *f* of .4 is generally considered large.

With *N* close to 20 for all groups, this ANOVA was relatively robust to violations of the assumption of normality. Students had no contact between grade level, and worked individually on the assessment, maintaining the assumption of independence. Based on the descriptive statistics in Table 2, there does not seem to be any substantial violation of the normality assumption (skewness < 1.0 for all groups) or of the equal variance assumption (maximum standard deviation ratio = 1.1).

An ANOVA was conducted examining the three different groups taking the test, a summary of which is on Table 3. The obtained $F(2,62) = 7.67, p = .0011$, was judged to

be statistically significant at the .05 level suggesting that at least one pair of population group means differ. A *Tukey* test determined that the 6th grade group differs significantly from the 10-12th grade group.

Table 3

One-Way ANOVA Summary Table of 6th, 8th, and 10-12th Grade

Source	<i>Df</i>	<i>MS</i>	<i>F</i>	<i>p</i>
Between	2	295.76	7.67	.0011
Within	62	38.57		

This pilot suggests there is evidence to contradict the theory that the learning/practice window closes at the age of ten.

Study Population and Sample

Assessment A, developed and refined from the previous two pilot studies, was administered to 292 students at a public, magnet school for visual and performing arts in North Carolina. Students participating in the study were enrolled in the instrumental music education program at the school. The test was administered during regular instrumental classes over a two-day period as a part of their daily class routine. Each class was given instructions and examples to practice on together as a group prior to their taking the test to insure they understood the task. Students marked their answers on a scantron sheet. In addition to the answers, students included their grade level, age, gender, and whether they had any private lesson experience on their papers. Scantron sheets were turned in without names on the paper to assure anonymity. Approval was granted by both the principal and the district before the test was administered.

Variables

The primary independent variable of interest in the study was the age of each subject. Gender of the subject and whether or not each student had any private music lesson experience was also considered.

The dependent variables in the study included the scores on the speech and music items of the assessment.

Measures

The measures utilized in the study were scores obtained from the test prepared to assess the listening skill along with responses to self-report information obtained at the time the test was administered.

Data Analysis

To address the questions posed in this study, three 7x2x2 factorial ANOVAs were performed examining the test scores of students with their age, gender, and prior private lesson experience. The first ANOVA examined the three independent variables with the overall test score of each subject. The other two ANOVAs examined the same three independent variables with each subjects' score of the music and speech portions of the test by themselves. Because three tests were being applied to the same data, a *Bonferroni* adjustment was used on the initial alpha level of .05 to account for the rising Type I error rate. The adjusted alpha level for the three tests was .017.

Chapter Four

Results

The purpose of this study was to investigate the effect age had on the ability of students older than the proposed learning window (11-18 years of age) to discriminate subtle changes in phrase interpretation. For analysis, students were divided initially into eight age groups: 11-year-olds ($n = 42$); 12-year-olds ($n = 57$); 13-year-olds ($n = 57$); 14-year-olds ($n = 48$); 15-year-olds ($n = 39$); 16-year-olds ($n = 29$); 17-year-olds ($n = 15$); and 18-year-olds ($n = 5$). The group of 18-year-olds with an n of 5 was too small to include in this investigation as its own group. It was combined with the 17-year-old group (17-18-year-old group $n = 20$).

Assumptions

Before three 7x2x2 level factorial ANOVAs, each with three between group factors, could be performed, three assumptions were determined to have been met in the study. The first was that the scores within groups could be considered independent because students did not interact during the treatment and were monitored by three teachers.

The second assumption was that the population distribution was normal. Each age group was evaluated in terms of mean, SD, skewness, and kurtosis (see Table 4). Generally accepted ranges for skewness and kurtosis are -1.0 to 1.0 and -0.5 to 1.0, respectively. Groups aged 11, 14, 15, and 16 displayed kurtosis scores which gave cause for further examination. In all groups mentioned, the distributions were slightly

platykurtic. The first factorial ANOVA examining the overall test scores was robust to violations of the assumption of normality because the sample sizes were large. The last assumption was that of the homogeneity of population variances. Based on the standard deviation statistics in Table 4, there did not seem to be any violation of the equal variance assumption.

Table 4

Age 11, 12, 13, 14, 15, 16, & 17-18 Overall Test Score ANOVA Descriptive Statistics

	Age Groups						
	11	12	13	14	15	16	17-18
<i>N</i>	42	57	57	48	39	29	20
Mean in %	56.55	58.43	64.25	62.40	61.73	69.48	63.13
S.D. in %	17.98	17.03	17.70	18.76	21.13	21.18	19.25
Skewness	-.17	-.27	-.54	-.40	-.18	-.52	-.73
Kurtosis	-1.24	-.22	-.42	-.53	-.93	-.90	-.14

Sample means of the groups found on Table 4 did not appear to have much variation. To provide a standardized measure of effect size, *Cohen's f* was calculated to be .18. This could be interpreted to mean that the group means deviate about .18 standard deviation units from the grand mean. According to Cohen, this indicated a small effect size. Because none of the assumptions had been violated, an ANOVA could be conducted.

In anticipation of running three factorial ANOVAs on the same data set to address the questions posed at the beginning of the study, the initial $\alpha = .05$ was adjusted down to

$\alpha = .017$ using a *Bonferroni* adjustment. This was necessary to account for an increasing Type I error rate as more than one test was run on the same data set.

A *Pearson Product-Moment correlation coefficient* of .75 was produced when examining students' scores on the music items with those of the speech items. This score indicates a high positive correlation. An r^2 value of .56 indicates that 56% of the total variance of music items is explained by the variance of the speech items. *Cronbach's alpha* on the overall sample was calculated at .87. According to Nunnally (1978) this score is within the established range of acceptability for internal reliability.

Overall Test Results

Once the assumptions were determined to have been met, the first 7x2x2 factorial ANOVA, three between group factors was performed. This ANOVA examined the independent variables of the seven different age groups, gender, and prior lesson experience with the dependant variable of overall test score to determine any possible levels of significance. A summary for the factorial ANOVA examining age, gender and lessons on the overall test is displayed in Table 5.

Table 5

7 x 2 x 2 Factorial ANOVA Summary Table Investigating the Relationship Between Age, Gender, and Prior Lesson Experience on Students' Ability to Interpret Music and Speech Phrase Changes

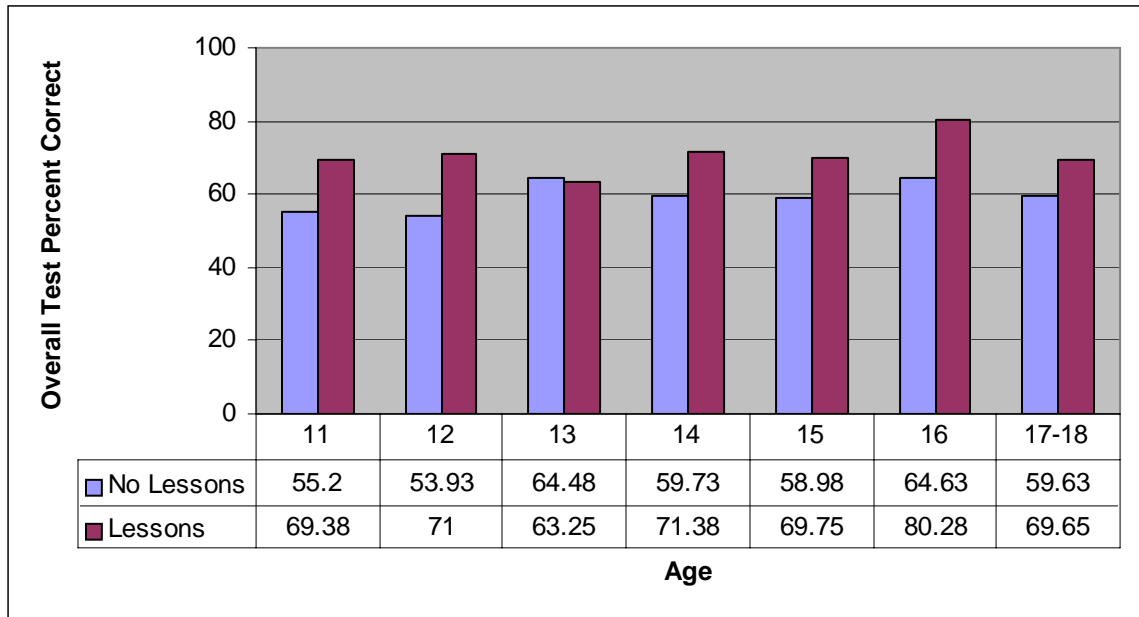
Source	<i>Df</i>	<i>MS</i>	<i>F</i>	<i>r</i> ²
Age	6	51.52	.98	
Gender	1	206.19	3.93	
Lesson	1	672.93	12.83*	.04
Age x Gender Int.	6	16.11	.31	
Age x Lesson Int.	6	38.96	.74	
Gender x Lesson Int.	1	12.71	.24	
Age x Gender x Lesson Int.	6	42.58	.81	
Within	264	52.46		

* Significant at $p < .017$ after Bonferroni adjustment for a study-wide α level of .05.

This overall analysis revealed a significant main effect for lessons, $F(1, 264) = 12.83$; $p < .0004$. Students with private lesson experience, $n = 66$, Mean Score = 70.72%, scored significantly higher than those students with no private lesson experience, $n = 226$, Mean Score = 59.2%. Sample means for lesson experience arranged by age group are displayed in Figure 1.

Figure 1

Overall Lesson Score Means by Age Group in Percentages



The main effect for age was nonsignificant, $F(6, 264) = .98; p = .44$. The main effect for gender was nonsignificant, $F(1, 264) = 3.93; p = .048$. The interaction of age and gender was nonsignificant, $F(6, 264) = .31; p = .93$. The interaction of age and lessons was nonsignificant, $F(6, 264) = .74; p = .62$. The interaction of gender and lessons was nonsignificant, $F(1, 264) = .24; p = .62$. The interaction of age, gender, and lessons also was nonsignificant, $F(1, 264) = .81; p = .56$.

Speech and Music Parts

An examination of the test parts, music and speech, using two 7x2x2 factorial ANOVAs with three between group factors was performed to determine any possible

levels of significance. Before the tests were conducted, descriptive statistics were examined for any anomalies that would have violated normality assumptions.

Descriptive statistics for speech items grouped by age are located in Table 6. All age groups fell within the accepted range of skewness and groups aged 11, 13, 14, 15, 16, and 17-18 demonstrated slightly platykurtic distributions. The large group sizes made the ANOVA robust to this normality violation. There did not seem to be any violation of the equal variances assumption.

Table 6

Age 11, 12, 13, 14, 15, 16, & 17-18 Speech Items ANOVA Descriptive Statistics

	Age Groups						
	11	12	13	14	15	16	17-18
<i>N</i>	42	57	57	48	39	29	20
Mean in %	61.45	62.55	67.80	66.15	66.30	71.55	67.25
S.D. in %	21.9	19.2	21.35	22.1	23.15	23.6	19.95
Skewness	-.25	-.53	-.47	-.62	-.33	-.71	-.42
Kurtosis	-1.24	-.12	-.56	-.58	-1.04	-.84	-.72

Descriptive statistics for music items grouped by age are located in Table 7. All age groups fell within the accepted range of skewness and groups aged 11, 12, 16, and 17-18 demonstrated slightly platykurtic distributions. The 17-18-year-old group demonstrated a notably leptokurtic distribution. The equal variance assumption had not been substantially violated.

Table 7

Age 11, 12, 13, 14, 15, 16, & 17-18 Music Items ANOVA Descriptive Statistics

	Age Groups						
	11	12	13	14	15	16	17-18
<i>N</i>	42	57	57	48	39	29	20
Mean in %	51.65	54.30	60.70	58.65	57.20	67.40	59.00
S.D. in %	17.40	17.60	16.55	17.85	21.25	20.8	20.15
Skewness	.03	.22	-.36	-.01	-.1	-.14	-.63
Kurtosis	-.7	-.57	-.2	-.34	-.42	-1.23	1.01

A summary for the factorial ANOVA examining age gender and lessons for speech items is displayed in Table 8.

Table 8

7 x 2 x 2 Factorial ANOVA Summary Table Investigating the Relationship Between Age, Gender, and Prior Lesson Experience on Students' Ability to Interpret Speech Phrase Changes

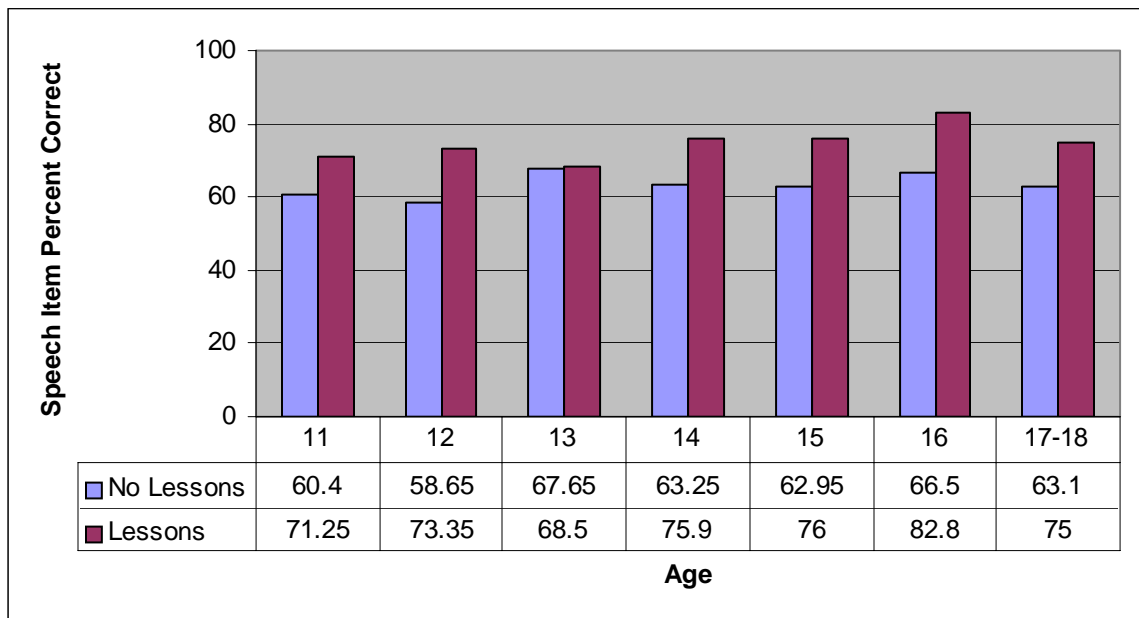
Source	<i>Df</i>	<i>MS</i>	<i>F</i>	<i>r</i> ²
Age	6	11.34	.64	
Gender	1	58.5	3.29	
Lesson	1	185.47	10.42*	.03
Age x Gender Int.	6	5.7	.32	
Age x Lesson Int.	6	6.67	.37	
Gender x Lesson Int.	1	.12	.01	
Age x Gender x Lesson Int.	6	12.21	.69	
Within	264	17.8		

* Significant at $p < .017$ after Bonferroni adjustment for a study-wide α level of .05.

The factorial ANOVA of the speech items revealed a significant main effect for lessons, $F(1, 264) = 10.42$; $p < .001$. Students with private lesson experience, $N = 66$, Mean Score = 74.75%, scored significantly higher on speech test items than those students with no private lesson experience, $N = 226$, Mean Score = 63.1%. Sample means for lesson experience arranged by age group are displayed in Figure 2.

Figure 2

Speech Item Score Means for Lesson Experience Arranged by Age Group in Percentages



The main effect for age was nonsignificant, $F(6, 264) = .64$; $p = .7$. The main effect for gender was nonsignificant, $F(1, 264) = 3.29$; $p = .07$. The interaction of age and gender was nonsignificant, $F(6, 264) = .32$; $p = .93$. The interaction of age and

lessons was nonsignificant, $F(6, 264) = .37; p = .89$. The interaction of gender and lessons was nonsignificant, $F(1, 264) = .01; p = .94$. The interaction of age, gender, and lessons also was nonsignificant, $F(1, 264) = .69; p = .66$.

A summary for the factorial ANOVA examining age, gender and lessons for music items is displayed in Table 9.

Table 9

7 x 2 x 2 Factorial ANOVA Summary Table Investigating the Relationship Between Age, Gender, and Prior Lesson Experience on Students' Ability to Interpret Music Phrase Changes

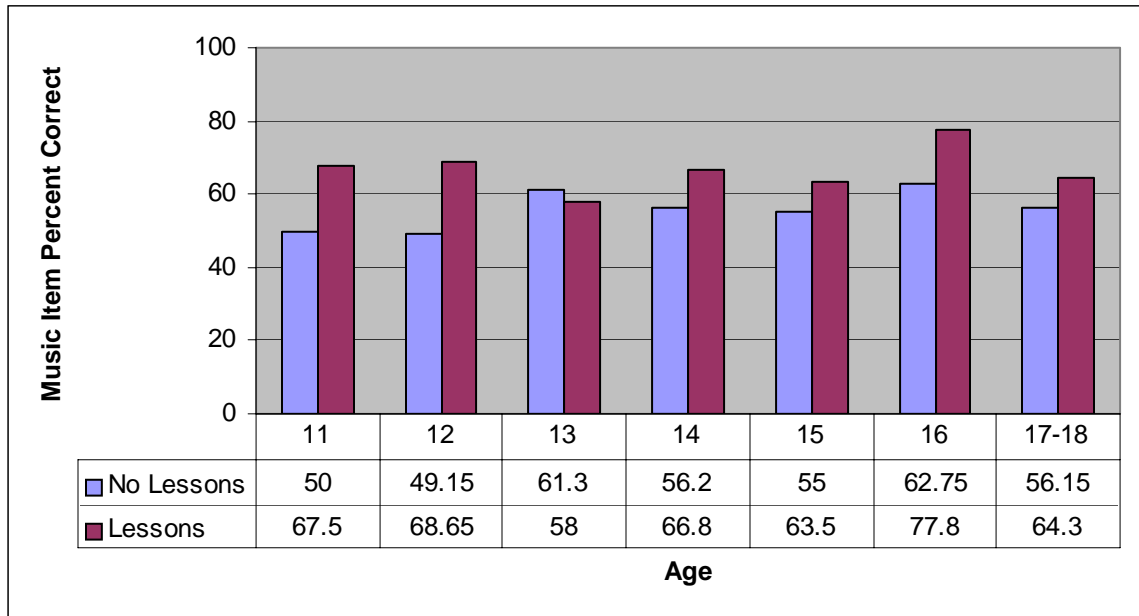
Source	<i>Df</i>	<i>MS</i>	<i>F</i>	<i>r</i> ²
Age	6	14.74	1.17	
Gender	1	45.03	3.57	
Lesson	1	151.84	12.04*	.04
Age x Gender Int.	6	6.21	.49	
Age x Lesson Int.	6	14.95	1.19	
Gender x Lesson Int.	1	10.38	.82	
Age x Gender x Lesson Int.	6	12.19	.97	
Within	264	12.61		

* Significant at $p < .017$ after Bonferroni adjustment for a study-wide *alpha* level of .05.

The factorial ANOVA of the music part of the test revealed a significant main effect for lessons, $F(1, 264) = 12.04; p < .0006$. Students with private lesson experience, $N = 66$, Mean Score = 66.65%, scored significantly higher on music test items than those students with no private lesson experience, $N = 226$, Mean Score = 55.35%. Sample means for lesson experience arranged by age group are displayed in Figure 3.

Figure 3

Music Item Score Means for Lesson Experience Arranged by Age Group in Percentages



The main effect for age was nonsignificant, $F(6, 264) = 1.17; p = .32$. The main effect for gender was nonsignificant, $F(1, 264) = 3.57; p = .06$. The interaction of age and gender was nonsignificant, $F(6, 264) = .49; p = .81$. The interaction of age and lessons was nonsignificant, $F(6, 264) = 1.19; p = .31$. The interaction of gender and lessons was nonsignificant, $F(1, 264) = .82; p = .36$. The interaction of age, gender, and lessons also was nonsignificant, $F(1, 264) = .97; p = .44$.

In an examination of all three factorial ANOVAs conducted for this study, the r^2 accounted for no more than 3-4% of the variance. This indicates that while there was a significant difference observed between those students who had participated in lessons

and those who had not, it only explains a very small portion of the total variance. It is also noteworthy that in all three tests conducted, no significance was found between age group and students' ability to discriminate changes in instrumental and speech phrase interpretations.

Chapter Five

Discussion

The purpose of this study was to determine what effect age had on middle school and high school students' ability to interpret subtle changes or nuances in music and speech phrases. It was designed to further test a proposed theory of a learning window for the perception of expressive qualities in music and speech. The theory suggests that a practice window must overlap a learning window before it closes around the age of 10 to properly develop music prosody and speech prosody skills.

This study is a continuation and advancement of the work accomplished by Heller and Athanasulis (2002). In their study, Heller and Athanasulis developed a test comprised of thirty speech items and thirty music items, each consisting of three small phrases. It was administered to forty 1st-grade students (about six years old), forty 3rd-grade students (about eight years old), and forty 5th-grade students (about 10 years old).

Using an analysis of variance, Heller and Athanasulis (2002) found that in music items, growth was found to be statistically significant between the 6- and 10-year-olds (10.7%). A *Scheffé's post hoc* comparison showed the mean difference between 6- and 8-year-olds was not significant (4.8%), but was significant in 8- to 10-year-olds (5.8%). This small increase showed no leveling off in subjects as a function of age, which contradicted the proposed theory.

In an analysis of the speech items, significant growth was found in both groups (17.3% in 6- to 8-year olds and 10.7% in 8- to 10-year olds). The researchers concluded

that the reduction in growth between the pairwise comparisons in the speech items suggests a leveling off of growth which supports the theory.

One limitation found in the Heller and Athanasulis study was in the development of the test items. An item analysis was not conducted, and it is possible that better items could have been selected given the large number of item possibilities to choose from. Another limitation was in the choice of students. While the goal of the study was to test a theory of a learning window that theoretically closes at the age of around ten, no students older than around ten were investigated. Closure of this learning window was not fully investigated.

The current study attempted to follow up on the work of Heller and Athanasulis by first improving on the quality of the test. Three hundred sixty items were developed and then tested to determine through item analysis which questions were most appropriate for the task. In addition, this study investigated the ability of students who were older than the proposed learning window theory to determine when and if the learning window closes at the proposed age of 10 to 11 for either music or speech phrase interpretation.

In all three factorial ANOVA tests of the entire test scores and their parts, no significance was found between any groups aged 11 to 18. This contradicts the results of the second pilot study which found that older students improved in ability to take the test. One possible explanation for this difference in the pilot study is the smaller sample sizes and different schools being examined between the sixth and eighth graders and the 10-12th grade group. The results of the current study reinforce the proposed theory that a learning/practice window seems to be in place for the acquisition of auditory

interpretative skill. What is of particular interest in this finding is that the results seem to contradict the natural refinement of the quality of music student over time. Younger groups (aged 11-12) are comprised of any student expressing interest in learning instrumental music. As years pass, students who remain in instrumental music show an increasing interest and longer-term commitment to the development of their musical craft. The reduction of the retention in the program can be seen in the sizes of the groups examined in this study. It is noteworthy that students who remained in instrumental music for six years of public school music education did not fare statistically better than those just entering the program. At the time of the testing, beginning sixth graders would have only had one semester of instrumental music education. These results seem to indicate a leveling off of students' ability to interpret speech and music phrase changes.

In all three factorial ANOVAs, prior lesson experience demonstrated the only level of significance in the three independent variables or their interactions. Those students who had indicated they participated in private lessons in the past scored significantly higher than those students who had not taken private lessons previously. A possible explanation for this outcome may be the additional time devoted to music study implied by their decision to pursue additional music education outside the public school classroom. Those students allocating time and money on supplemental music education may have a more vested interest in the improvement of their musical abilities. Additional time practicing and studying music is inferred with the additional lessons. It is interesting to note that those students who enrolled in private music lessons in the past also scored higher than those students who did not take lessons on the speech portion of the test. This

may be due to a higher overall commitment to academic excellence in general on the part of those students opting for additional music lessons.

Limitations

A number of factors could have affected the data obtained in this study. While the selection of a school containing sixth to twelfth grade students selected by lottery from the city population provided a good representation, a more useful sample may have been random subjects taken from a variety of different schools to better represent the general music education environment. While the students obtained consistent education from the same music teachers throughout their tenure at the school, their experience and the results of this study may not be generalizable to the rest of the instrumental music education environments.

Another limitation was in the wording of the private lesson question. Students were asked if they had ever participated in private lessons in the past. The question was not specific to private lessons prior to the closing of the proposed learning window. This means that a sixteen-year-old student could have answered yes to the private lesson question, but had lessons at the age of 14, well after the proposed learning window would have closed. While the private lesson information is valuable, it could have been more useful with a more specifically worded question.

The length of the test may have been too long for consistent concentration and focus of each student on the task. Perhaps providing a break in the middle of the test would have improved students' level of focus. In addition, it is possible that smaller groups may have provided a quieter environment to better hear the subtle changes in interpretation.

Future Study

The results of this study reinforce the idea that a practice/learning window may exist for the development of the speech and music interpretation skill. If the theory is correct and a practice/learning window does seem to be in place for the acquisition of an interpretative skill in music and speech, music educators must reexamine the order in which their pedagogy falls. Some forms of music deemed valuable and important for public appreciation by the profession, with the exception of sporadic good fortune in the occasional primary music education classroom, are not formally introduced until well after the proposed learning/practice window closes. Jazz studies are routinely omitted from music education until late in secondary education, if at all. Masterworks throughout history are often not introduced until secondary education. Comprehensive world music study is sometimes neglected until college, even for music majors.

One reason for this is that the genres called to attention are more technically advanced than others. More technically challenging music is put off until students have an ability to successfully accomplish performing the music. In some cases this means late in their secondary education.

If there is validity to this theory, exposure to these kinds of music is occurring too little and too late in students' development. In fact, it could be argued that one reason students have a more difficult time negotiating these forms of music could be because they have had so little experience with it earlier in their education. This is not an advocacy for starting brass quintets and jazz bands in primary school. Instead, it may be a call for a more comprehensive, in depth and experiential learning program for children within the learning/practice window to the music of history and the world.

Future studies are needed to further explore the proposed learning/window theory. An examination of students who had private lesson experience within the learning window may demonstrate that more extensive music activity within the learning window may lead to improved ability to interpret phrase changes.

A study using current popular or world music phrases to examine people's ability to choose which one of three phrases is most culturally appropriate should be considered. If young participants are able to easily determine the appropriate phrase but older participants cannot, it may reinforce the suggested theory.

It is important to continue to research and develop a better understanding as to how the rules of music listening are learned. Understanding this process can lead to more efficient, age-appropriate pedagogy in music education.

References

- Atterbury, B. W. (1983) A comparison of rhythm pattern perception and performance in normal and learning-disabled readers, age seven and eight. *Journal of Research in Music Education*, 31 (4), 259-270.
- Blacking, J. (1973). *How musical is man?* Seattle: University of Washington Press.
- Boyle, J. D. & Radocy, R. E. (1987). *Measurement and evaluation of musical experiences*. New York: Schirmer Books.
- Campbell, M. R. (1991). Musical learning and the development of psychological processes in perception and cognition. *Bulletin of the Council for Research in Music Education*, 107, 35-48.
- Cox, S. M. O. (1977). *A descriptive analysis of the responses to beat, meter, and rhythm pattern by children, grades one to six*. The University of Wisconsin-Madison, Ph.D.
- Davies, J. B. (1978). *The psychology of music*. California: Stanford University Press.
- Davidson, L. & Scripp, L. (1985). Young children's musical representations: windows on music cognition. In J. A. Sloboda (Ed.), *Generative Processes in Music : The Psychology of Performance, Improvisation, and Composition* (pp. 195-230). New York, NY: Oxford University Press.
- Denardo, G. F., & Kantorski, V. J. (1995). A Continuous Response Assessment of Children's Music Cognition. *Bulletin of the Council for Research in Music Education*, 126, 42-52.

- De Yarmen, R. (1975). An investigation of the stability of musical aptitude among primary-age children. *Experimental Research in the Psychology of Music, Studies in the Psychology of Music*, 10. Iowa City: University of Iowa Press.
- Elliott, D. J. (1995). *Music matters: A new philosophy of music education*. New York: Oxford University Press.
- Fiske, H. (1990). *Music and mind: philosophical essays on the cognition and meaning of music*. New York: Edwin Mellen Press.
- Fiske, H. (1996). *Selected theories of music perception*. New York: Edwin Mellen Press.
- Flavell, J. H. (1965). *The developmental psychology of Jean Piaget*. New York: D. Van Nostrand Company.
- Flohr, J. W. (1981). Short-term music instruction and young children's developmental music aptitude. *Journal of Research in Music Education*, 29 (3), 219-224.
- Gardner, H. (1983). *Frames of mind: The theory of multiple intelligences*. New York: BasicBooks.
- Gibson, B. L. (1986). *Young children's response to interpretation in music and speech*. University of Connecticut, Ph.D.
- Gordon, E. E., *Advanced Measures of Music Audiation*, Chicago, IL: G.I.A. Publications, Inc., 1989.
- Heller, J. (2003). *The ability to categorize interpretations of music and speech phrases*. a paper presented at the Research Alliance in Music Education Meeting, Corsham, England.
- Heller, J. and Athanasulis, M. (2002). Music and language: A learning window. *Bulletin of the Council for Research in Music Education*, 153, 18-22.

- Heller, J. J. (2000). An implicit rule system for music listening. *Bulletin of the Council for Research in Music Education, 147*, 77-80.
- Heller, J. and Campbell, W. (1981). A theoretical model of music perception and talent. *Council for Research in Music Education, 66-67*, 20-24.
- Heller, J., Campbell, W., and Gibson, B. (1982). The development of music listening skills in children. *Psychology of Music, 10*, (3), 55-58.
- Heller, J. & Campbell, W. (1982). Music communication and cognition. *Council for Research in Music Education, 72*, 1-15.
- Holahan, J. M., Saunders, T. C., & Goldberg, M. D. (2000). Tonal Cognition in Pattern Discrimination: Evidence from Three Populations. *Journal of Research in Music Education, 48*(2), 162-177.
- Hufstader, R. A. (1977). An investigation of a learning sequence of music listening skills. *Journal of Research in Music Education, 25*, 185-196.
- Johnson, C.M. (2000). Effect of adding interpretive elements to a musical performance on the rhythmic and dynamic variations. *Bulletin of the Council for Research in Music Education, 147*, 91-96.
- Jones, R. (1976). The development of the child's conception of meter in music. *Journal of Research in Music Education, 24* (3), 142-154.
- Krumhansl, M. (1983). Perceptual structures for tonal music. *Music Perception, 1*(1), 28-59.
- Lerdahl, F., & Jackendoff, R. (1983). *Theories of developmental psychology*. San Francisco: W. H. Freeman.
- Levman, B.G. (1992). The genesis of music and language. *Ethnomusicology, 36*, 147-70.

- London, J. (1996). Musical and linguistic speech acts. *The Journal of Aesthetics and Art Criticism*, 54, 49-64.
- McBeth, W. F. (1992). Interpretation: unlocking the drama in music. *The Instrumentalist*, 47, 14-18.
- Meyer, L. B. (1956). *Emotion and Meaning in Music*. Chicago: University of Chicago Press.
- Miller, L. K., and Eargle, A. (1990). The contributions of development versus music training to simple tempo discrimination. *Journal of Research in Music Education*, 38 (4), 294-301.
- Moles, A. (1958). *Information theory and esthetic perception*. Urbana: University of Illinois Press.
- Nash, M. (1997, February 3). Fertile minds. *Time*. 48-56.
- Nettl, B. (1983). *The Study of Ethnomusicology*. Urbana: University of Illinois Press.
- Nettl, B. (1994). "Musical thinking" and "thinking about music" in ethnomusicology: an essay of personal interpretation [Blackfoot, Persian, classical]. *The Journal of Aesthetics and Art Criticism*, 52, 139-48.
- Nunnally, J. C. (1978). *Psychometric theory*. 2nd ed. New York: McGraw Hill.
- Palmer, M. (1976). Relative effectiveness of two approaches to rhythm reading for fourth-grade students. *Journal of Research in Music Education*, 24 (3), 110-118.
- Perney, J. (1976). Musical tasks related to the development of the conservation of metric time. *Journal of Research in Music Education*, 24 (4), 159-168.

- Pflederer, M. (1967). Conservation laws applied to the development of musical intelligence. *Journal of Research in Music Education*, 15(3), 215-223.
- Rodriguez, C.X. and Webster, P. R. (1997). Development of children's verbal interpretative responses to music listening. *Bulletin of the Council for Research in Music Education*, 134, 9-30.
- Schleuter, S. and DeYarmen, R. (1977). Musical aptitude stability among primary school children. *Council for Research in Music Education*, 51, 14-22.
- Seashore, C. (1938). *Psychology of music*. New York: Dover Publications,
- Seashore, C. (1919). *The psychology of musical talent*. Boston: Silver, Burdett and Company.
- Serafine, M. L. (1983). Cognition in Music. *Cognition*, 14, 119-183.
- Serafine, M.L. (1988). *Music as cognition: The development of thought in sound*. New York: Columbia University Press.
- Shaffer, L. H. & Todd, N. (1994). The interpretive component in musical performance. In R. Aiello (Ed.), *Musical Perceptions* (258-272). New York: Oxford University Press.
- Shehan, P. K. (1987). Effects of Rote Versus Note Presentations on Rhythm Learning and Retention. *Journal of Research in Music Education*, 35 (2), 117-126.
- Sloboda, J. A. (1985). *The musical mind: The cognitive psychology of music*. Oxford: Clarendon Press.
- Swain, J.P. (1996). The range of musical semantics. *The Journal of Aesthetics and Art Criticism*, 54, 135-52.

Zimmerman, M. P. (1982). Developmental processes in music learning. In R. Colwell (Ed.), *Symposium in Music Education: A Festschrift for Charles Leonard*.

Urbana: University of Illinois Press.

Zimmerman, M. P. (1986). Music development in middle childhood: A summary of selected studies. *Bulletin of the Council for Research in Music Education*, 86, 18-35.

Zimmerman, M. P. (1993). An Overview of Developmental Research in Music. *Bulletin of the Council for Research in Music Education*, 116, 1-21.

Appendices

Appendix A

Percentage of Students who Answered Correctly Each Item of the Five Speech Categories. Each category's items are unrelated to the other categories listed.

Item	Speech Items				
	Child Speech Only	Female Speech Only	Male Speech Only	Child, Female Male Speech	Female, Child Male Speech
1	95	76	100	83	68
2	95	100	100	83	84
3	68	53	100	78	53
4	95	88	89	89	42
5	82	59	83	83	84
6	95	94	94	22	63
7	0	65	61	39	21
8	55	88	72	83	89
9	41	41	83	78	32
10	14	71	72	28	89
11	82	100	89	83	26
12	41	47	89	44	53
13	05	41	67	67	63
14	64	0	72	72	79
15	68	71	67	61	32
16	23	53	17	39	58
17	95	100	94	89	32
18	59	88	50	67	42
Items					
Selected:	8, 9	3, 12	7, 18	12, 16	3, 12

Appendix B

Percentage of Students who Answered Correctly Each Item of the Five Nonsense Categories. Each category's items are unrelated to the other categories listed.

Nonsense Speech Items					
Item	Child Only	Female Only	Male Only	Female, Child Male	Child, Female Male
1	100	52	74	09	49
2	87	92	95	65	71
3	0	72	63	70	41
4	67	68	89	52	41
5	80	56	79	35	76
6	13	52	68	57	24
7	67	60	42	52	59
8	87	64	53	48	35
9	13	24	84	43	29
10	93	32	32	22	35
11	47	68	53	70	47
12	53	76	79	17	12
13	07	72	37	74	76
14	100	68	84	65	71
15	87	56	74	57	53
16	100	28	42	39	35
17	07	16	79	48	59
18	0	60	68	91	41
Items					
Selected:	11, 12	1, 6	8, 11	4, 17	11, 15

Appendix C

Percentage of Students who Answered Correctly Each Item of the Five Musical Except Categories. Each category's items are unrelated to the other categories listed.

Musical Except Instrumental Items						
Item	Clarinet Only	Cello Only	Piano Only	Piano, Cello Clarinet	Cello, Clarinet Piano	
1	92	38	60	12	13	
2	63	31	65	41	17	
3	50	100	45	59	39	
4	63	88	30	88	22	
5	38	75	45	29	48	
6	79	38	30	35	30	
7	75	88	35	59	61	
8	71	69	50	35	43	
9	71	50	40	76	35	
10	71	94	55	35	30	
11	58	50	55	35	30	
12	58	50	45	29	57	
13	67	75	45	47	39	
14	71	44	45	29	35	
15	42	50	50	29	17	
16	79	69	60	29	57	
17	33	88	45	35	57	
18	63	31	50	35	57	
Items						
Selected:	3, 12	9, 11	8, 18	3, 13	5, 17	

Appendix D

Percentage of Students who Answered Correctly Each Item of the Five Composed Melodies Categories. Each category's items are unrelated to the other categories listed.

Composed Melody Instrumental Items					
Item	Clarinet Only	Cello Only	Piano Only	Piano, Cello Clarinet	Cello, Clarinet Piano
1	56	83	62	26	38
2	72	61	29	65	44
3	50	48	43	61	38
4	78	96	14	09	50
5	67	61	76	74	50
6	78	65	67	35	50
7	67	61	62	52	56
8	67	17	67	52	63
9	50	87	67	43	75
10	83	96	62	43	63
11	72	22	67	30	56
12	72	78	14	48	56
13	56	78	81	35	38
14	67	70	38	26	88
15	83	74	38	26	69
16	78	83	57	35	88
17	67	65	33	52	31
18	50	0	19	78	56
Items					
Selected:	3, 9	2, 3	3, 16	8, 12	4, 5

Appendix E

Pilot One: Descriptive Statistics

Speech					
	N	Mean/Total	% Correct	Median	SD
Child	22	10.8/18	60	10.7	2.5
Female	18	11.7/18	65	11.6	3.3
Male	18	14/18	77.78	14.5	2.8
Child, Female, Male	20	10.7/18	59.44	11.6	4.2
Female, Child, Male	19	10.1/18	56.1	9.5	2.6
Nonsense					
Child	15	10.1/18	56.1	10.5	2.1
Female	25	10.2/18	56.67	9.6	3.7
Male	19	11.9/18	66.11	11.5	2
Child, Female, Male	17	8.4/18	46.67	7.7	2.5
Female, Child, Male	23	9.1/18	50.56	9.5	3.5
Melodic Excerpt					
Clarinet	20	8.5/18	47.22	7.7	2.6
Cello	24	11.4/18	63.33	11.6	3.3
Piano	16	11.3/18	62.78	10.6	2.9
Piano, Cello, Clarinet	17	7.4/18	41.11	7.7	1.8
Clarinet, Cello, Piano	23	6.9/18	38.33	7.6	2.5
Composed Melody					
Clarinet	21	9/18	50	8.7	2.5
Cello	18	12.1/18	67.22	11.6	3
Piano	23	11.6/18	64.44	11.6	3.3
Piano, Cello, Clarinet	23	7.9/18	43.89	7.7	2.6
Clarinet, Cello, Piano	16	10.1/18	56.1	10.6	3

About the Author

Andrew Sioberg received a Bachelor's Degree in Music from Radford University in 1994, a Master's Degree in Jazz Studies and Contemporary Media from the Florida State University in 1997, and a Master's Degree in Music Education from Syracuse University in 1999. Following his training at Syracuse University, Andrew taught instrumental music education at the Durham School of the Arts in Durham, North Carolina for three years. He entered the Ph.D. program at the University of South Florida in 2002 and completed the degree in 2005. Throughout his education and teaching, Dr. Sioberg maintained an active career as a professional performer, arranger, and clinician.