TOWSON UNIVERSITY

OFFICE OF GRADUATE STUDIES

TEST-RETEST RELIABILITY OF THREE TESTS OF TEMPORAL PROCESSING IN NORMAL HEARING YOUNG ADULTS

By:

Crystal Lilly, B.S.

A Thesis

Presented to the faculty of

Towson University

in partial fulfillment

of the requirement for the degree

Doctor of Audiology

Department of Audiology, Speech-Language Pathology and Deaf Studies

Towson University
Towson, Maryland 21252

May 2014

APROVAL PAGE

TOWSON UNIVERSITY
OFFICE OF GRADUATE STUDIES

THESIS APPROVAL PAGE

This is to certify that the thesis prepared by Crystal Lilly, B.S. entitled Test retest reliability of three tests of temporal processing in normal hearing adults has been approved by thesis committee as satisfactorily completing the thesis requirements for the degree Doctorate of Audiology.

Dr. Stephanie Nagle, Ph.D.	10/28/13 Date
Dr. Jennifer Smart, Ph.D.	11/4/13 Date
Dr. Shannon Palmer, Ph.D.	10/26/2013 Date
Committee Member	Date
Committee Member	Date
Committee Member	Date
Dean of Graduate Studies	Date

ACKNOWLEDGMENTS

I would like to thank so many people as I would not have been able to complete my thesis without their guidance, support and direction. First, I would like to thank my thesis chair, Dr. Stephanie Nagle, Ph.D., and committee, Dr. Jennifer Smart, Ph.D., and Dr. Shannon Palmer, Ph.D., for their continued guidance and direction throughout the thesis process. I would also like to thank my colleague and peer, Lindsay Sperling, B.S., for her continued support. I'm not sure I would have made it through this program without you by my side pushing me through the tough times and reminding me to laugh through it all. Additionally, I would like to thank my friends and family for the support they have given me throughout my lifetime, I would not be where I am today without all of you.

ABSTRACT

There is a lack of reliability data within the literature for behavioral tests of auditory processing disorder (APD). Thus, the purpose of this study was to evaluate the test-retest reliability of three tests of temporal processing (Frequency Pattern, Duration Pattern, and Gaps In Noise tests) in normal hearing, young adults. The methods of this study included administering the Frequency Pattern (FP), Duration Pattern (DP) and Gaps In Noise (GIN) tests according to each test's owner's manual to normal hearing adults at one test session and then again seven to nine days later. The data was analyzed using a general linear model analysis of variance (ANOVA) for each test individually in order to compare the mean scores obtained at test session one to the mean scores obtained at test session two. Additionally, Pearson *r* correlation coefficients were analyzed for each test in order to investigate the reliability of each test. The results of this study indicated that there were no significant differences between the mean scores of any of the tests administered across testing sessions. Additionally, moderate to strong correlation data was found ($.446 < r < .71$). Thus, this study concludes that the FP, DP, and GIN tests are reliable clinical tools.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

KEY TO ABBREVIATIONS

AAA: American Academy of Audiology

ABR: Auditory brainstem response

ADHD: Attention Deficit Hyperactivity Disorder

ANOVA: Analyses of Variance

ALR: Auditory late response

APD: Auditory Processing Disorder

ASHA: American Speech-Language Hearing Association

CANS: Central auditory nervous system

CD: Compact disk

CNS: Central nervous system

daPa: daca-pascal

dB: Decibel

dB HL : Decibel hearing level

dB SL: Decibel sensation level

DDD: Dichotic double digits

DP: Duration Pattern

FP: Frequency Pattern

GIN: Gaps In Noise

Hz: Hertz

IRB: Institution Review Board

MLD: Masking level difference

MLR: Middle latency response

MMN: Mismatch negativity

Msec: millisecond

PPT: Pitch Pattern Test

PTA: Pure tone average

SCAN-3 A: SCAN version 3 for Adolescents and Adults

SCAN-3 C: SCAN version 3 for Children

SCAN A : SCAN for Adolescents and Adults

SCAN C: Scan for Children

SD: Standard deviations

SLI: Specific Language Impairment

SRT: Speech recognition threshold

T1: Test session one

T2: Test session two

TBIs: Traumatic brain injuries

CHAPTER 1

Introduction

APD is defined as a deficit or impairment in an individual's ability to process auditory information at the level of the central auditory nervous system (CANS) despite normal peripheral hearing structures (ASHA, 2005a; Jerger & Musiek, 2000; Lovett & Johnson, 2010; Neijenhuis, Snik, & Broek, 2003). It has been recommended and discussed in the literature that tests for APD should be sensitive, specific, and reliable and studies reporting this information are currently lacking within the peer-reviewed literature (AAA, 2010; ASHA, 2005a; Jerger & Musiek, 2000; Strouse & Hall, 1995). An accurate diagnosis of Auditory Processing Disorder (APD) is essential for rehabilitation purposes.

Currently, there is no standard test battery used for the diagnosis of APD. A high degree of sensitivity and specificity are essential elements of tests that make up a test battery; however, test-retest reliability data is also an important factor to consider when compiling a test battery for APD (AAA, 2010; ASHA, 2005a; Jerger & Musiek, 2000). APD is typically diagnosed based on behavioral tests and due to the inherent subjectivity of behavioral testing, test-retest reliability data is essential (Strouse & Hall, 1995). This data is essential because if tests used during the diagnosis of APD are reliable, then any observed change following intervention therapy may be attributed to an improvement in performance rather than to test-retest reliability errors (Theunissen, Swanepoel, & Hanekom, 2009; Wilson & McArdle, 2007). However, test-retest reliability data on commercially available behavioral tests utilized for the diagnosis of APD is lacking (AAA, 2010; Strouse & Hall, 1995).

As mentioned previously, there is no standard test battery for APD; however, inclusion of temporal processing tests within the test battery has repeatedly been recommended throughout the peer-reviewed literature (AAA, 2010; ASHA, 2005b; Bellis, 2003, 2004; Chermak & Musiek, 1997; Cox, McCoy, Tun, & Wingfield, 2008). Several commercially available temporal processing tests have proven to be sensitive and specific; however, test-retest reliability data for these tests is absent within the literature (AAA, 2010; Musiek, Baran, & Pinheiro, 1990; Musiek & Pinheiro, 1987; Musiek et al., 2005; Musiek et al., 2011). Thus, the purpose of this study is to identify the test-retest reliability of three tests of temporal processing that have been found to be sensitive and specific (Frequency Pattern, Duration Pattern and Gaps-In-Noise) in normal hearing individuals in order to validate the use of these tests within an APD test battery.

Chapter 2

Literature Review

Auditory processing is a complex, systematic process in which a signal arriving at an individual's ear is detected, transmitted, filtered, and in most cases, perceived and understood as having meaning. As defined by the American Speech-Language Hearing Association (ASHA) technical report, auditory processing is broadly referred to as the "efficiency and effectiveness by which the central nervous system utilizes auditory information" (2005a, p. 2). The understanding and comprehension of auditory information is a much more complex process than mere detection of an auditory signal. Detection, identification, and comprehension of an auditory stimulus undergoes a multifaceted process which involves serial and parallel processing within many anatomical and neural mechanisms of the central nervous system (CNS) (American Academy of Audiology [AAA], 2010). A lot of progress has been made in the identification of the central function of the auditory system and the neural mechanisms which are involved in normal auditory processing (Kaga, Shindo, Tanaka, & Haebara, 2000; Musiek & Baran, 2004; Musiek, Baran, & Pinheiro, 1990; Musiek et al., 2007; Musiek, Charette, Morse, & Baran, 2004). However, when there is a disconnect, or when these auditory structures and/or neural mechanisms function abnormally and present as listening difficulties, this is known as an auditory processing disorder (APD) and can manifest in a variety of ways.

APD is defined as a deficit or impairment in "the perceptual processing of auditory information in the CNS" (ASHA, 2005a p. 2). APD is characterized by a deficit

in the ability to process, or interpret, audible sounds which are transmitted to the central auditory nervous system (CANS) with normal peripheral hearing structures (Jerger & Musiek, 2000; Lovett & Johnson, 2010; Neijenhuis, Snik, & Broek, 2003). Individuals diagnosed with APD may have difficulty with auditory processes such as (a) sound lateralization and localization; (b) auditory discrimination; (c) auditory pattern recognition; (d) temporal processing skills including temporal integration, discrimination, ordering, and masking; (e) auditory abilities with degraded signals; and/or (f) auditory abilities with competing signals (ASHA, 2005a; Beck & Bellis, 2010; Jerger & Musiek, 2000; Musiek & Chermak, 1994). Symptoms of a deficit in these auditory processing skills include, but are not limited to, difficulty with (a) understanding speech in the presence of background noise, (b) concentrating in noisy situations, (c) following multistep oral directions, (d) maintaining attention, and/or (e) understanding rapid or degraded speech (AAA, 2010; Bamiou, Musiek, & Luxon, 2001; Beck & Bellis, 2010; Dawes, Bishop, Sirimanna, & Bamiou, 2008; Jerger & Musiek, 2000; Neijenhuis et al., 2003). However, due to the heterogeneous organization of the CANS and the complexity in which auditory processing occurs, symptomology and behaviors of individuals with APD may vary from patient to patient (Jerger & Musiek, 2000). Additionally, it is important to note that APD is a deficit in the processing of auditory stimuli that is not due to any other higher order impairments such as language, learning, or other cognitive factors (ASHA, 2005a).

In addition to diverse symptomology of APD, the etiology of APD is not always known and can be displayed in a variety of populations. Individuals with known deficits that affect the CANS, as well as those with neuro-degenerative diseases (e.g. Alzheimer's

disease, etc.) and traumatic brain injuries (TBIs) may have an APD (AAA, 2010; Chedru, Bastard, & Efron, 1978; Colson, Robin, & Luschei, 1991; Gates, Anderson, Feeney, McCurry, & Larson, 2008; Grimes, Grady, Foster, Sunderland, & Patronas, 1985; Rance et al., 2010; Strouse, Hall, & Burger, 1995; Turgeon, Champoux, Lepore, Leclerc, & Ellemberg, 2011). In addition, individuals with neurodevelopmental disorders and communicative disorders such as language disorders, learning disabilities, and attention deficit hyperactivity disorder (ADHD) have been reported in the literature as having auditory processing difficulties co-occurring with their other disorder(s) (AAA, 2010, Cook et al., 1993; Ferre & Wilber, 1986, Huang et al., 2012; Riccio, Hynd, Cohen, Hall, & Molt, 1994; Witton, 2010). However, individuals with no other difficulties can also exhibit auditory processing problems and, often, the etiology is unknown (AAA, 2010; Middelweerd, Festen, & Plomp, 1990; Neijenhuis, Stollman, Snik, & Van den Broek, 2001).

Comorbidity and APD

The brain does not function in homogenous "compartments" and the brain's neural pathways and mechanisms have a heterogeneous organization which results in complex processing schemes (Alho et al., 1996). Similarly, the CANS processes information via serial and parallel processing schemes. Due to the complexity of the CANS, and the heterogeneous population that is affected by APD, individuals with APD may exhibit other difficulties in addition to impairments in auditory processing abilities (ASHA, 2005; Bamiou et al., 2001). Several studies have found that APD can occur in conjunction with difficulties in higher order functioning such as learning, speech, language, and other related processes (ASHA, 2005a; Bamiou et al., 2001; Bellis &

Ferre, 1999; Chermak & Musiek, 1997; Jerger & Musiek, 2000; King, Lombardino, Crandell, & Leonard, 2003; Sharma, Purdy, & Kelly, 2009; Witton, 2010).

King et al. (2003) examined the comorbidity of adults with developmental dyslexia and APD. These researchers assessed the temporal processing abilities in adults aged 21 to 32 years with dyslexia as compared to age and intelligence-matched peers. The temporal processing abilities were assessed as part of the study due to the hypothesis that various language and reading disabilities (such as Specific Language Impairment [SLI] and dyslexia) stem from an impairment in the individual's temporal processing abilities which may disrupt the individual's normal acquisition of critical language skills (King et al., 2003; Tallal, 1980; Tallal & Piercy, 1973). Thus, King et al. (2003) used the Frequency Pattern test (Pinheiro & Ptacek, 1971), Duration Pattern test (Musiek et al., 1990), and a gap detection test (Tucker-Davis Technologies system II) to assess the temporal processing abilities of adult students (mean age = 24.4 years) with compensated developmental dyslexia. These researchers found that as a group, the adults with compensated developmental dyslexia performed significantly poorer ($p \leq 0.002$) than age and intelligence-matched controls on the Frequency Pattern and Duration Pattern tests; however, no difference between the groups was found for the gap detection test (King et al., 2003). These results support the need for a valid, comprehensive, multidisciplinary evaluation for individuals with various difficulties in order to provide appropriate intervention and rehabilitation therapies based on accurate diagnoses as APD can be a comorbid disorder.

To add to the literature on comorbidity and APD, Sharma et al. (2009) conducted a comprehensive study which assessed a range of auditory, language and reading abilities

in 68 children aged 7-12 years with either suspected or confirmed APD to determine the percentage of children with "pure" APD and children with APD and coexisting language and/or reading difficulties. The researchers for this study utilized five behavioral tests for auditory processing, three tests for reading ability, and one test to evaluate cognition, language, short term memory, and sustained attention skills. Results from this study found that 47% (n= 32) of the children assessed were diagnosed with APD, a language impairment, and a reading disorder, whereas only 4% (n= 3) were diagnosed with APD alone. Sharma et al. (2009) hypothesized that the overlap found between APD and other disorders is due to the lack of assessment tools available that adequately distinguish between auditory, reading and language dysfunction. Thus, this again supports the notion that the assessment of individuals suspected of APD and other higher-order disorders should reflect a multidisciplinary approach and use valid tests with high sensitivity, specificity and test-retest reliability in order to accurately identify each patient's specific area of weaknesses, or difficulties (AAA, 2010; ASHA, 2005a; Cacace & McFarland, 1998; Jerger & Musiek, 2000; Sharma et al., 2009).

While studies have found a high percentage of comorbidity among APD and other disorders, it should be noted that the relationship between APD and other higher order difficulties (such as language) is complex and does not illustrate a one-to-one relationship (AAA, 2010; ASHA, 2005a; Cacace & McFarland, 1998). As stated previously, the etiology of APD is often unknown and various combinations of auditory deficits can present as a variety of functional difficulties (Jerger & Musiek, 2000). Due to this heterogeneity and complexity of APD, the associated impact APD has on an individual and their higher order processes may vary according to the degree of neurologic

involvement as well as other factors such as social and environmental influences (ASHA, 2005a). Thus, it is important that a valid, comprehensive, multidisciplinary team approach is utilized for the assessment of individuals with suspected APD in order to accurately diagnosis APD and/or other disorders and enable the professional(s) to identify specific areas of difficulties and in turn, implement appropriate intervention programs (Bamiou et al., 2001).

APD Tests

There is no standard diagnostic test battery for APD; however, due to the complexity with which the brain processes auditory stimuli, the test battery for APD should not only assess auditory processing skills, it should also assess the integrity of the various auditory structures at different levels of the CANS (AAA, 2010; ASHA, 2005b; Bamiou et al., 2001; Bellis, 2003, 2004; Chermak & Musiek, 1997; Cox, McCoy, Tun, & Wingfield, 2008; Neijenhuis et al., 2003). Therefore, a diagnosis of APD should only be made after a comprehensive auditory processing assessment has been performed. This comprehensive test battery should, ideally, include sensitive and specific behavioral and electrophysiologic tests coupled with an in-depth case history (AAA, 2010; Bamiou et al., 2001; Jerger & Musiek, 2000).

It has been suggested and generally agreed upon, that after normal outer, middle and inner ear functions have been established, the behavioral APD test battery should include a variety of different tests. Variety of tests is important in order to assess different levels of the CANS (ASHA, 2005a). These tests may include: (a) dichotic listening tasks, (b) monaural low-redundancy speech tests, (c) tests of temporal

processing and (d) binaural interaction tests (Bellis, 2004; Bellis & Ferre, 1999; Chermak & Musiek, 1997; Jerger & Musiek, 2000).

Dichotic listing tasks are important to include as part of a test battery as they assess an individual's ability to separate, or integrate, competing stimuli presented to each ear simultaneously (ASHA, 2005a; Noffsinger, Martinez, & Wilson, 1994). Dichotic speech tests have a variety of linguistic test materials which can include digits, syllables, words, or sentences (AAA, 2010; Noffsinger et al., 1994). Additionally, certain dichotic tests have been proven to be highly sensitive to dysfunction at the level of the auditory cortex (AAA, 2010; Musiek, 1983a; Meyers, Roberts, Bayless, Volkert, & Evitts, 2002).

Monaural low-redundancy speech tests assess an individual's ability to recognize acoustically or digitally degraded speech stimuli which are presented to the participant monaurally (one ear at a time) (ASHA, 2005a; Bellis, 2003). Tests of monaural low-redundancy measures an individual's auditory closure and discrimination skills when speech is presented in a filtered, or distorted, fashion or when it is embedded within background noise (Bellis, 2003). Thus, these tests have been shown to be useful in describing the auditory abilities of individuals; however, they are less sensitive in identification of APD as compared to other behavioral tests and are more vulnerable to being affected by language and cognitive abilities (AAA, 2010; Musiek & Baran, 2002; Musiek, Chermak, Weihing, Zappulla, & Nagle, 2011).

Temporal processing tests assess an individual's ability to recognize and sequence auditory stimulus patterns, and evaluate acoustic stimuli over time (ASHA, 2005a; Bellis,

2003). These skills are important for speech perception and awareness of rhyming patterns (Chermak & Lee, 2005; Rawool, 2007). Temporal processing abilities can be broken up into different sub-categories which include temporal patterning, temporal resolution (or gap detection), temporal integration, and temporal masking (ASHA, 2005a). Additionally, several temporal processing tests, including temporal patterning and temporal gap detection tests, have proven to have a high degree of sensitivity and specificity in identifying individuals with APD with confirmed lesions and have wide clinical utility (AAA, 2010; Musiek & Pinheiro, 1987; Musiek et al., 2005; Musiek et al., 2011).

Binaural interaction tests are useful for the assessment of an individual's localization and lateralization skills. Localization and lateralization skills are dependent on the binaural evaluation of intensity or timing differences of acoustic stimuli between ears (ASHA, 2005a). Tests of binaural interaction, such as the masking level difference (MLD) test, has proven to be sensitive to APD as a consequence to lower-level brainstem dysfunction (AAA, 2010; Lynn, Gilroy, Taylor, & Leiser, 1981); however, commercial tests of binaural interaction which are valid and efficient for clinical practice are minimal (AAA, 2010; Musiek et al., 2011).

Electrophysiological tests are also recommended as part of a comprehensive test battery for the diagnosis of APD in order to objectively measure the functioning of anatomical structures from the level of the brainstem up to the cortex (AAA, 2012; Jerger & Musiek, 2000). Unlike behavioral tests, electrophysiologic tests are a useful tool in the identification of APD as they are not confounded by other variables such as cognitive and intellectual disabilities, attention deficit disorders, and/or speech and language

delays/impairments (Jerger & Musiek, 2000). Additionally, electrophysiologic responses can be elicited with a variety of stimuli from simple tone bursts to complex speech signals (AAA, 2010). Some electrophysiologic tests that are recommended to support the identification of APD are the auditory brainstem response (ABR), middle latency response (MLR), the auditory late response (ALR), mismatch negativity (MMN), and P300 (AAA, 2010; Bamiou et al., 2001; Jerger & Musiek, 2000).

The literature has shown that various electrophysiologic tests are a valuable tool for the assessment of the functioning of various anatomical structures (AAA, 2010; Baran, Bothfeld, & Musiek, 2004; Hall & Johnston, 2007; Hillyard, Hink, Schwent, & Picton; 1973; Jerger et al., 1991; Jirsa & Clontz, 1990; Musiek, Charette, Kelly, Lee, & Musiek, 1999; Musiek & Lee, 1995; Tremblay, Kraus, McGee, Ponton, & Otis, 2001). However, there are several clinical issues to consider prior to administering any electrophysiologic tests as part of an APD test battery (AAA, 2010). A potential downfall is the fact that they take a considerable amount of time to administer and the clinician must weigh the cost-effectiveness of these measures (Jerger & Musiek, 2000). In addition, there is no global set of protocols for stimulus parameters and measurement and there is little to no normative data for individuals across the lifespan as many tests have a high degree of inter-subject variability (AAA, 2010). Additionally, a survey of clinicians by Emanuel, Ficca, and Korczak (2011) revealed that electrophysiologic testing for APD is not widely used as only 7.7% always did an electrophysiologic test as a part of their APD test battery out of 372 respondents. Due to the variety of confounding factors impacting the use of electrophysiologic tests in the APD test battery, as well as

the fact that APD is typically diagnosed based on behavioral testing, behavioral tests will be the main focus of this discussion.

Although there is no widely used test battery for the diagnosis of APD, there is consensus among professionals and clinicians for the diagnostic criteria required for a diagnosis of APD, which is a failure of at least two standard deviations (SD) below the mean on two different behavioral tests or more than three SD below the mean on one behavioral test in at least one ear (AAA, 2010; ASHA, 2005a; Bellis, 2003; Chermak & Musiek, 1997). The performance on the tests administered should reflect the areas of dysfunction and auditory difficulties, and guide appropriate intervention therapies (AAA, 2010). Thus, a test should be valid and a failure on a specific test of APD should reflect a dysfunction, or difficulty, in the area of auditory processing that is assessed (i.e. tests of temporal processing should actually assess an individual's temporal processing abilities so that in turn, a failure on a test of temporal processing reflects a weakness in the individual's temporal processing abilities). Again, as discussed, it is essential that a valid and reliable test battery be administered in order to accurately diagnose an individual with APD and guide intervention therapies.

Diagnostic accuracy of tests for APD. The selection of tests which comprise an APD test battery should have a high degree of validity, reliability and efficiency (sensitivity and specificity) (AAA, 2010; ASHA, 2005a; Bellis, 2003; Chermak & Musiek, 1997). The validity of a test is the extent to which the test measures what it is intended to measure; in the case of APD, a valid test should accurately measure the central auditory processing abilities of an individual (Ostergard, 1983; Theunissen, Swanepoel, & Hanekom, 2009). Reliability of a test is the extent to which the test's

results remain consistent across different testing times (Ostergard, 1983; Theunissen et al., 2009). To add to this, reliability of a test's results with multiple lists must also be equivalent between test lists (Nilsson, Soli, & Sullivan, 1994; Ostergard, 1983; Theunissen et al., 2009). It has been reported that a test has adequate test-retest reliability when the measure has a strong correlation coefficient (r) of 0.80 or more and no significant differences between the means (Amos & Humes, 1998; Groth-Marnat, 2009). Sensitivity is the ability of a test to correctly identify affected individuals as having a dysfunction/pathology and is calculated by dividing the number of correctly identified individuals with a known lesion/pathology by the number of individuals with a lesion that were assessed (Musiek et al., 2011; Theunissen et al., 2009). On the other hand, specificity of a test is the test's ability to correctly identify unaffected individuals as having normal functioning and is calculated by dividing the number of correctly identified individuals as normal functioning by the total number of normal functioning participants tested (Musiek et al., 2011; Theunissen et al., 2009). Tests which have been found to have the above characteristics will have a high degree of diagnostic value and are essential clinical tools in the diagnosis of APD (Musiek et al., 2011).

Sensitivity and specificity are important to consider when developing a test battery for the evaluation of APD to ensure diagnostic efficiency (AAA, 2010; ASHA, 2005a; Musiek et al., 2011). Data from commercially available behavioral tests which have reported information on sensitivity and specificity are outlined in Table 1. From a review of the literature, the Frequency Pattern and Duration Pattern tests have the greatest degree of sensitivity and specificity (Musiek et al., 1990; Musiek et al., 2011; Musiek & Geurkink, 1982; Musiek & Pinheiro, 1987). This agrees well with the results from a

study conducted by Musiek et al. (2011) which evaluated the diagnostic accuracy of the Dichotic Double Digits, Competing Sentences, Frequency pattern and Low-Pass Filtered Speech tests for a control group (n= 29, mean age= 27.0 years and SD = 10.5) and a group with normal hearing despite confirmed lesions to the CANS (n=20, mean age = 28.7 years, SD = 12.2) and found that the Frequency Pattern test had the highest overall sensitivity, specificity, and efficiency.

Table 1.

Reported Sensitivity and Specificity of Commercially Available Behavioral Tests of Auditory Processing

Test	Sensitivity	Specificity	Reference
Dichotic			
Dichotic Double Digits	75-90%	83-91%	Hurley & Musiek, 1997; Musiek, 1983a; Musiek et al., 2011
Dichotic Three Digit Form	50%	75%	Mueller, Beck, & Sedge, 1987
Competing Sentences	25-75%	100%	Domitz & Schow, 2000; Musiek, 1983b; Musiek et al., 2011
Monaural Low Redundancy			
Low Pass Filter Speech	50-75%	-	Karlsson & Rosenhall,1995; Lynn & Gilroy, 1977; Musiek et al., 2011
Temporal Processing			
Frequency Pattern	83-90%	88-95%	Musiek et al., 2011; Musiek & Geurkink, 1982; Musiek & Pinheiro, 1987
Duration Pattern	86%	92%	Musiek et al., 1990
Pitch Pattern	30%	100%	Domitz & Schow, 2000
Gaps In Noise	67%	94%	Musiek et al., 2005

Note. Review of the literature on the sensitivity and specificity of commercially available tests is presented above.
- Not reported.

Test-retest reliability. Test-retest reliability is an essential factor to consider during the selection of tests to include in an APD test battery. Test-retest reliability is especially important to consider when selecting tests to measure improvement in auditory

skills following aural rehabilitation. The purpose of investigating the reliability of a test is to determine the degree of variability of a test which is caused by error of the test (Groth-Marnat, 2009). Variability of a test across test sessions is inevitable, especially since some variability occurs due to true fluctuations in an individual's performance between test sessions; this is true variability that is unrelated to test error (Groth-Marnat, 2009). Although it is impossible to control for the natural variability in an individual's score due to the nature of human performance, a goal of test construction is to keep testing errors to a minimum and to design a test in such a way that reduces variability that is a function of the test itself (Groth-Marnat, 2009). Test-retest correlation coefficients indicate the extent to which the scores obtained across test sessions can be generalized from one situation to the next and are calculated by correlating the scores obtained across test sessions from the same individuals (Groth-Marnat, 2009). If the degree of correlation between the scores is high, than the clinician can assume that the scores obtained accurately reflect the performance of an individual and changes in score are less likely to be due to random fluctuations in the test (Groth-Marnat, 2009; Theunissen et al., 2009; Wilson & McArdle, 2007). A high correlation, and thus, high test-retest reliability is considered to be .80 or better; however, it has been stated that when utilizing tests which are used to make decisions about individuals (tests which diagnose disorders), clinicians should use tests which have correlations of .90 or better (Amos & Humes, 1998; Groth-Marnat, 2009). Conversely, Polite and Beck (2012) indicated that for quantifying the degree of correlation for subjective tests, like the ones used for the diagnosis for APD, Spearman correlation coefficient values of 0.7-1.0 are considered high correlations, 0.41-0.69 are considered moderate correlations, 0.2-0.4 are considered

low correlations and anything below 0.2 are considered to have no statistical correlation. Due to the nature of the testing and the subjective measures used in the current study, the above cutoff criterion will be used for the evaluation of the statistical data obtained in the current study.

In terms of auditory processing tests, when APD is diagnosed via tests with high validity, by definition, an area of weakness or difficulty will be reflected on the failed auditory tests in order to drive intervention therapies (AAA, 2010). If tests used to assess APD are also reliable, then any observed change following intervention therapy may be attributed to a true improvement in performance rather than to test-retest reliability errors (Groth-Marnat, 2009; Theunissen et al., 2009; Wilson & McArdle, 2007). Test-retest reliability data of behavioral tests is critical due to the nature of these tests and the fact that they are subjective tests. Behavioral testing has inherent variability due to the nature of human performance (Groth-Marnat, 2009; Strouse & Hall, 1995). Thus, since APD is typically diagnosed via behavioral tests, high test-retest reliability for these tests is critical. Additionally, when using tests with higher linguistic loads or multiple test lists, learning effects and inter-list equivalency must be considered as they contribute to the reliability of a test measure (Groth-Marnat, 2009; Theunissen et al., 2009).

For tests of auditory processing, reliability data within the literature for commercially available tests is outlined in Table 2. However, test-retest reliability data on other commercially available tests utilized for the diagnosis of APD is lacking (AAA, 2010; Strouse & Hall, 1995).

Table 2.

Reported Test-Retest Reliability Data of Behavioral Tests of Auditory Processing

Test/Subtest	Significant difference between the means?	Correlation Coefficients	Between Test Time	Reference
Dichotic Double Digits	No	.77 < r < .97	2 months – 1 year	Musiek et al., 1991; Strouse & Hall; 1995
Gaps In Noise	No	.88 < r < .95	At least 1 week	Musiek et al., 2005
Pitch Pattern 4 item	Yes $p < .05$	r = .91	7-10 days	Summers, 2003
Pitch Pattern 3 item	No	.65 < r < .99	1-2 weeks; 4 months	Domitz & Schow, 2000; Humes, Coughlin, & Talley, 1996
Duration Pattern 4 item	No	r = .90	7-10 days	Summers, 2003
Duration Pattern 3 item	No	r = .80	4 months	Humes et al., 1996
Competing Sentences	Yes	.57 < r < .82	1 – 2 weeks	Domitz & Schow, 2000; Summers, 2003
SCAN	--	.22 < r <.75	1 -2 weeks; 6-7 weeks; 6 months	Amos & Humes, 1998; Keith, 1986
SCAN A	---	.5 < r < .74	1 day – 5 months	Keith, 1995; Spencer, 2007
SCAN C	--	.65 < r < .82	--	Keith, 2000
SCAN-3 A	--	.54 < r < .80	1 – 29 days	Keith, 2012a; Lovett & Johnson, 2010
SCAN-3 C	--	.54 < r < .73	--	Keith, 2012b

Note. Review of the literature on the reliability of behavioral tests of auditory processing is presented above. SCAN A = SCAN for Adolescents and Adults. SCAN C = SCAN for Children. SCAN-3 A = SCAN version 3 for Adolescents and Adults. SCAN-3 C = SCAN version 3 for Children

- Not reported.

Temporal Processing

Temporal processing refers to the ability of the brain to process and interpret time-related aspects of acoustic signals (ASHA, 2005a). Temporal processing skills include temporal ordering (temporal sequencing), temporal resolution (gap detection), temporal integration, and temporal masking (forward and backward masking) (ASHA, 2005a). Individuals with temporal processing difficulties report difficulty understanding speech in the presence of background noise, difficulty understanding speech when multiple speakers are present, and/or difficulty understanding fast-talking speakers (Baran et al., 2004). These complaints are often the most reported complaints among all individuals with APD and the elderly population (Chermak & Musiek, 1997; Dawes et al., 2008; Fitzgibbons & Gordon-Salant, 1996; Martin & Jerger, 2005). To add to this, there are several studies throughout the literature that suggest temporal processing abilities play a role in the identification of critical components of speech; thus, it has been hypothesized that underlying deficits in temporal processing may result in difficulty understanding speech as well as have an influence on co-morbid disorders related to speech and language and/or reading disabilities (Bellis, 2003; Buonomano & Karmarkar, 2002; Fitzgibbons & Gordon-Salant, 1996; Helfer & Vargo, 2009; Houtgast & Festen, 2008; Martin & Jerger, 2005; Musiek & Chermak, 1994; Wright, Buonomano, Mahncke, & Merzenich, 1997). Thus, temporal processing tests are a critical tool in the evaluation of individuals suspected of having APD.

Timing-related cues, and in turn, temporal processing, are key to understanding and formulating speech sounds (Musiek, Shinn, & Hare, 2002). Intact temporal processing abilities are needed in order to comprehend speech as phoneme

differentiation, syllabic rhythm, and perception of pitch related to varying rates (or timing differences) of the vocal fold vibration (Phillips, 2002). Additionally, prosodic cues within speech (such as pauses and duration of speech sounds) which are detected via temporal processes provide semantic information (Samelli & Schochat, 2008). Due to the fact that intact temporal processing is needed in order to effectively understand speech, several investigators have hypothesized about the correlation between language based disorders (including reading, dyslexia, phonological awareness, Specific Language Impairment [SLI]) and temporal processing difficulties (King et al., 2003; McArthur & Bishop, 2004; Merzenich et al., 1996; Tallal, 1980; Tallal & Piercy, 1973; Walker, Shinn, Cranford, Givens, & Holbert, 2002). However, the literature on this subject is varied and as mentioned previously, APD and its associated influence on other higher order difficulties, such as language, are complex and does not illustrate a one-to-one relationship (AAA, 2010; ASHA, 2005a; Cacace & McFarland, 1998).

Congruently, several studies have suggested that auditory difficulties experienced by the elderly population are the result of changes in central auditory processing which are unrelated to peripheral hearing loss (Chisolm, Willott, & Lister, 2003; Fitzgibbons & Gordon-Salant, 1996; Frisina & Walton, 2006; Jerger, Jerger, Oliver, & Pirozzollo, 1989; Stach, Loiselle, & Jerger, 1991). One of these changes includes a range of deficits with temporal processing abilities. This has been investigated by measuring the temporal processing abilities of young versus older populations. It has been found that the older populations score significantly poorer on tests of temporal processing as compared to younger populations (Grose, Hall, & Buss, 2006; Humes, Kewley-Port, Fogerty, & Kinney, 2010; Schneider, Pichora-Fuller, Kowalchuk, & Lamb, 1994; Snell, 1997;

Strouse, Ashmead, Ohde, & Granthan, 1998). However, one confounding factor in the evaluation of the elderly population for temporal processing deficits is the high likelihood of peripheral hearing impairment due to age (Kumar & Sangamanatha, 2011). Further research is needed on the effect of peripheral hearing impairment on tests of temporal processing.

Temporal processing tests are a critical part of the multidisciplinary evaluation of individuals with suspected APDs, reading and language disorders, and the elderly population in order for accurate differential diagnosis and recommendations for intervention therapies to take place. However, in order for an accurate diagnosis to be made, all tests used within a test battery should have a high degree of validity, accuracy and reliability; this reliability data is currently lacking within the literature for commercially available tests for APD (AAA, 2010; ASHA, 2005a; Bellis, 2003; Chermak & Musiek, 1997; Strouse & Hall, 1995). Thus, the purpose of this study is to identify the test-retest reliability of three tests of temporal processing (Frequency Pattern, Duration Pattern, and GIN) in normal hearing individuals in order to validate the use of these tests within an APD test battery.

Chapter 3

Methodology

Participants

Thirty normal hearing young adults between the ages of 18 and 30 were recruited for this study. Prior to testing, Institution Review Board (IRB) approval (Appendix A) was obtained and all participants signed an informed consent (Appendix B) and completed a comprehensive case history form (Appendix C). To qualify as a participant in this study, each participant exhibited pure-tone air-conduction thresholds < 25 dB HL for the octave frequencies between 250 and 8000 Hz and type A tympanograms which were defined as a static compliance of 0.3ml -1.4 ml, peak pressure within -150 daPa to + 100 daPa, and an ear canal volume of 0.6 ml -1.5 ml and/or symmetric volumes within .03 ml (Jerger, 1970).

Procedures

All participants were tested during two testing periods which were seven to nine days apart. Participants were tested while seated in a double-walled sound-attenuated booth. At the first testing period, otoscopic inspection, pure-tone air conduction testing and speech recognition threshold (SRT) testing was completed using the Grason Stadler (GSI) 61 two-channel diagnostic audiometer and calibrated EARTONE 3A insert earphones. Pure-tone testing was completed using the Modified Hughson-Westlake procedure (Carhart & Jerger, 1959) across the octave bands of 250 Hz to 8000 Hz and the participants were provided a push button to respond to the stimuli. Tympanometry was completed using a 226 Hz probe tone and the Madsen Otoflex 100 Middle Ear Analyzer. Both the GSI 61 audiometer and Madsen Otoflex Middle Ear Analyzer were calibrated

diagnostically on August 23, 2012 and were checked biologically prior to testing. Additionally at the first testing period, the auditory processing tests were administered in accordance with published, recommended procedures for each test. Procedures for these tests will be discussed in detail below. At the second testing period, otoscopic examination, tympanometry, and a threshold screening at 5 dB above each participant's thresholds (unless the participant's threshold was 25 dB HL at which time the threshold screening was at the threshold of 25 dB HL) that were obtained at the first testing session was completed. This was completed in order to ensure no conductive component had developed within the seven to nine days of between-testing time. After no change in middle ear integrity was confirmed, the auditory processing tests were re-administered using the same procedures as the first testing session.

Auditory processing test procedures and materials. All auditory processing test stimuli used for this study were prerecorded and routed through a Sony compact disc player to the GSI-61 two-channel audiometer and EARTONE 3A insert earphones. The Frequency Pattern (FP) and Duration Pattern (DP) test stimuli were prerecorded on the Audiology Illustrated compact disc (CD) and the Gaps-in-Noise (GIN) test stimuli were prerecorded on the GIN test CD. The order in which the auditory processing tests were administered and which ear they were presented to first was randomized and counterbalanced. The stimuli were calibrated to the audiometer prior to testing via a calibration tone which was adjusted to peak at 0 on the VU meter of the audiometer.

Frequency Pattern test. For the FP test, participants listened to the Audiology Illustrated pre-recorded CD which had 60 randomized test items. The test items consisted of a triad of 150 msec tone bursts (10 msec rise-fall times) which varied in

frequency between a low tone (880 Hz) or a high tone (1122 Hz). On this track, the stimuli were pre-recorded with an interstimulus interval of 200 msec with an approximately 6 second inter-item (or interpattern) interval. Each participant was required to verbally label the three-tone pitch pattern. There was a possibility of six combinations (e.g., high-high-low, high-low-high, low-high-high, low-low-high, low-high-low, high-low-low).

The FP test was administered at 50 dB HL as recommended by the test developer (Musiek, 2002). All participants were administered 3 practice items and 30 test items to each ear. The FP test was administered and instructions/training for the participant was completed in accordance with published recommended guidelines (Musiek, 2002). The participants were instructed that they would hear a pattern of three tones which vary in pitch and were required to verbally label the frequency pattern by stating "high-high-low", "low-high-low", etc. The participants were also instructed to guess if they were unsure of the pattern. The three practice test items were used to train the participant to the task, and were selected at random (a random starting point on the track). If the participant was still unsure of the task after three practice test items were completed, the participant was reinstructed and the same three practice items were re-administered. Following this training period, 30 test items were presented to each ear. This process was then completed to each ear again with seven to nine days between test dates. Note, the first 30 and second 30 test items were administered and the order of presentation and which ear (left versus right) received the test stimuli was randomized across the two testing sessions. Percent correct was calculated and reversals were not considered correct (Musiek, Pinheiro, & Wilson, 1980).

Duration Pattern test. For the DP test, participants listened to the Audiology Illustrated pre-recorded CD which had 60 randomized test items. The test items consisted of a triad of 1000 Hz tone bursts which varied in duration: short tones (250 msec) or long tones (500 msec). On this track, the stimuli were pre-recorded with an interstimulus interval of 300 msec with an approximately 6 second inter-item (or interpattern) interval. The participant was required to verbally label the three-tone pattern presented (a possibility of six combinations, e.g. long-long-short, long-short-long, short-long-long, short-short-long, short-long-short, long-short-short).

Like the FP test, the DP test was administered at 50 dB HL (Musiek et al., 1980). The participants were instructed and trained that they would hear a pattern of three tones which vary in duration. The participants were also instructed to guess if they were unsure of the pattern. Three practice test items were used to train the participant to the task, and were selected at random (a random starting point on the track). If the participant was still unsure of the task after three practice test items were completed, the participant was reinstructed and the same three practice items were administered. Following this training period, 30 test items were presented to each ear. Note, the first 30 and second 30 test items were administered and the order of presentation and which ear (left versus right) that received the test stimuli was randomized across the two testing sessions. This was then completed on each ear a second time with a seven to nine day inter-test time interval. Percent correct was calculated and reversals were not considered correct (Musiek et al., 1980).

GIN test. The GIN was used to evaluate the participant's ability to detect a gap of silence which was within a six-second duration of white noise. There were four different

lists which are equivalent in difficulty that each contained up to 36 different white noise presentations (Musiek et al., 2005). Within each noise presentation (or stimulus), there were zero to three gaps of silence. The gaps of silence varied in duration and would either be 2, 3, 4, 5, 6, 8, 10, 12, 15, or 20 msec long. Each of the 10 durations of gap silences appeared six times in random order within each of the four GIN lists. Thus, each list had a total of 60 test items with a silent 5 second interstimulus interval between test item (or noise segment). For this test, participants were required to press a response button when they heard a gap of silence within the white noise and were provided instructions as recommended within the literature (Musiek et al., 2005).

The GIN test was administered to each participant at 50 dB SL referencing each participant's pure tone average (PTA) (Musiek et al., 2005). Each participant was provided with the same instructions which were in accordance with published recommended guidelines (Musiek et al., 2005). Each participant was administered the practice test provided by the GIN. If the participant had difficulty with the practice items, the participant was reinstructed and the practice list was re-administered. Following the practice testing, one of the four test lists was administered to each ear in a randomized order. The GIN was scored by calculating the total number of correct for all gap durations by the total number of gap segments presented (i.e. x/60; see Table 3 for an example of a scored score-sheet) (Musiek et al., 2005). Additionally, the approximated gap detection threshold, which is defined as the shortest gap perceived by the participant at least 66.6% of the time, was calculated. However, if the participant obtained at least a 66.6% gap detection threshold at one gap duration but their performance on longer gaps was worse, the smallest gap that had at least a score of 66.6% consistently within the

longer gap durations was considered to be the gap detection threshold (Musiek et al., 2005).

Table 3.

Example of Gaps In Noise Scored Score Sheet

Gaps	2 msec	3 msec	4 msec	5 msec	6 msec	8 msec	12 msec	15 msec	20 msec	Total % Score
List 1	0/6	2/6	3/6	**5/6**	6/6	6/6	6/6	6/6	6/6	40/60
Left Ear	0%	33%	50%	**83%**	100%	100%	100%	100%	100%	66%

Note: Example of scoring sheet with the gap detection threshold bolded.

Data Analysis

Following data collection, the difference between test scores between test time one and test time two for the FP, DP, and GIN was analyzed using the IBM SPSS Statistics version 19. Descriptive statistics, a 2 tailed paired t-test, one way Analyses of Variance (ANOVA), general linear model repeated measures ANOVA, and Pearson *r* correlations were obtained. Statistical significance was determined utilizing an alpha level less than or equal to 0.05 ($p \leq 0.05$). Additionally, correlation coefficients (r values) of 0.7-1.0 were considered high, 0.41-0.69 were considered moderate, .02-.04 were considered poor and any r values below 0.2 were considered to have no statistical correlation (Polite & Beck, 2012). Individual data were also evaluated for clinical significance, which was any change in test score over the two testing periods which would qualify the individual to "pass" or "fail" the test and in turn, alter the diagnosis of APD.

CHAPTER 4

Results

Thirty adult participants were recruited for this study; however, three participants were lost due to attrition and one participant was excluded from data analysis due to a neurologic disorder. Thus, 26 participants (13 females and 13 males) were included in this study for data analyses. Participants ranged in age from 20 to 29 years (M = 23.6 years with SD 2.4 years). All participants were native speakers of English. Additionally, no participants had a history and/or suspicion of having an APD or any learning and/or language disorders. All participants had normal hearing and Jerger type A tympanograms bilaterally at both test session one (T1) and test session two (T2).

Auditory Processing Test Results

Descriptive statistics. Table 4 reports the mean scores for each test for left and right ears by test session. As mentioned previously, scores were reported as percent correct for the FP and DP test and as the gap detection threshold in milliseconds (msec) for the GIN test. Overall, the mean score on the FP test for both testing sessions ranged from 95.61% to 97.67% and 89.33% to 91.12% for the DP test. The mean scores for the GIN test ranged from 4.69 to 5 msec over both testing sessions.

Table 4.

Mean Scores for Left and Right Ears on Each Test for Test Session One and Two.

	Frequency Pattern		**Duration Pattern**		**Gaps In Noise**	
	Test 1	**Test 2**	**Test 1**	**Test 2**	**Test 1**	**Test 2**
Ear	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
Left	96.00 (6.93)	96.78 (5.93)	90.22 (9.48)	89.33 (12.00)	5 (1.13)	4.96 (0.82)
Right	95.61 (7.48)	97.67 (5.07)	90.35 (9.95)	91.12 (8.74)	5 (1.23)	4.69 (0.88)

Note. Above are the means and standard deviations for left and right ear scores for each test and test session The mean Frequency Pattern and Duration Pattern test scores and standard deviations are reported as percentages and the mean Gaps In Noise test scores and standard deviations are reported in milliseconds. SD = standard deviation and is reported in parentheses. Test 1 = test session 1. Test 2 = test session 2.

Ear effects. Prior to further data analyses, a 2 tailed paired t-test was completed to investigate ear effects on the FP, DP, and GIN test at T1 and T2. No significant difference between the means of scores between left and right ears was found for T1 or T2 for the FP ($p = .792$ and $p = .273$ respectively), DP ($p = .933$ and $p = .240$ respectively), and GIN test ($p = 1.00$ and $p = .283$ respectively). The 2 tailed paired t-test results can be seen in Table 5. A comparison between the means of left and right ears for each test can be visualized in Figure 1 and 2. Since no significant difference between ears was found for any of the tests, the data for each ear was collapsed for each test time for all subsequent analyses.

Table 5.

Paired 2 tailed T-Test Outcome for the Evaluation of Ear Effects for Each Test per Test Session

	Paired T-Test Outcome		
Test/Test Session	t	df	Sig. (2-tailed)
Frequency Pattern			
Test Session 1	3.420	25	.792
Test Session 2	-1.121	25	.273
Duration Pattern			
Test Session 1	-.084	25	.933
Test Session 2	-1.203	25	.240
Gaps In Noise			
Test Session 1	.000	25	1.00
Test Session 2	1.098	25	.283

Note. Above is the paired t-test outcome for the effect of ear stimulated for the Frequency Pattern, Duration Pattern, and Gaps In Noise test for each test session. t = t value. df = degrees of freedom. Sig. (2-tailed) = p value.

Figure 1. Mean left and right ear scores for the Frequency Pattern and Duration Pattern Test by test session

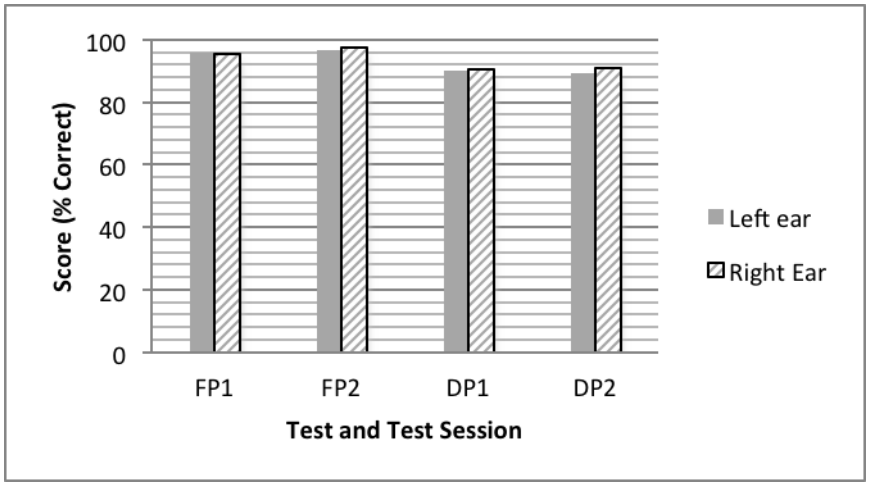

Figure 1. A graphical representation of the mean left and right ear scores for the Frequency Pattern and Duration Pattern test for test session 1 and test session 2. Mean score for left ear is denoted by solid fill and mean score for the right ear denoted by patterned fill. FP1 = Frequency Pattern test session 1. FP2 = Frequency Pattern test session 2. DP1= Duration Pattern test session 1. DP2 = Duration Pattern test session 2.

Figure 2. Mean left ear and right ear scores for the Gaps In Noise test by test session

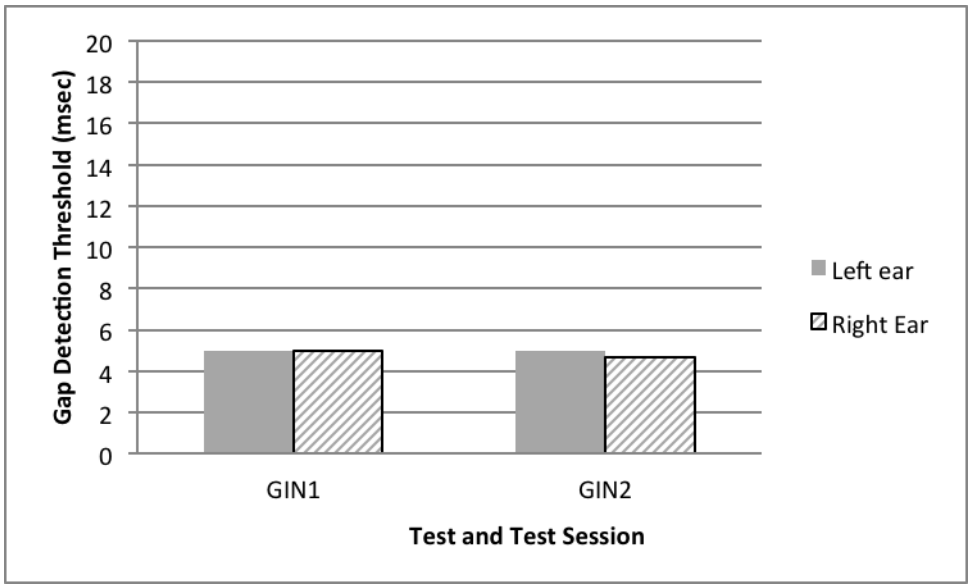

Figure 2. A graphical representation of the mean scores for each ear for the Gaps In Noise test for test session 1 and test session 2. The left ear mean score is denoted by solid fill and the right ear mean score is denoted by patterned fill. GIN1 = Gaps In Noise test session 1. GIN2 = Gaps In Noise test session 2.

Gender effects. The mean scores on the FP, DP, and GIN for males and females per test session can be found in Table 6. A one way ANOVA with a Bonferroni correction for the FP, DP and GIN test was completed to compare the means of male and female scores for each test session. Results indicated that there was no significant difference between the means of males and females for either test session except during T2 for the FP ($p = .021$) and GIN ($p = .034$) tests. The ANOVA outcomes can be seen in Table 7. A comparison between the mean score of males and females per test and test session is illustrated with significant differences denoted with an asterisk in Figure 3. Due to a significant difference in the mean of female and male scores for T2 for the FP

and GIN test, gender was used as a between subject factor for the subsequent repeated measures ANOVAs for all of the tests.

Table 6.

Mean Male and Female Scores by Test and Test Session with Standard Deviation

	Frequency Pattern		**Duration Pattern**		**Gaps In Noise**	
	Test 1	Test 2	Test 1	Test 2	Test 1	Test 2
Gender	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
Male	95.36 (7.85)	95.50 (7.18)	88.95 (11.78)	87.53 (13.28)	4.80 (1.16)	4.57 (0.57)
Female	96.26 (6.49)	98.95 (1.85)	91.63 (6.81)	92.92 (5.52)	5.19 (1.16)	5.07 (1.01)

Note. Above are the mean and standard deviations for male and female scores per test and test session. SD = standard deviation.

Table 7.

One Way ANOVA Outcomes for Effect of Gender for the Frequency Pattern, Duration Pattern, and Gaps In Noise Test for each Test Session

	One Way ANOVA Outcomes	
Test/Test Session	F	Sig.
Frequency Pattern		
Test Session 1	.203	.654
Test Session 2	5.64	.021*
Duration Pattern		
Test Session 1	1.01	.319
Test Session 2	3.65	.062
Gaps In Noise Test		
Test Session 1	1.41	.240
Test Session 2	4.75	.034*

Note. Above are the one way ANOVA outcomes for the effect of gender on the each test by test session. F= F statistic. Sig.= p value.
*= $p < .05$

Figure 3. Mean score of males and females for the Frequency Pattern and Duration Pattern Test by test session

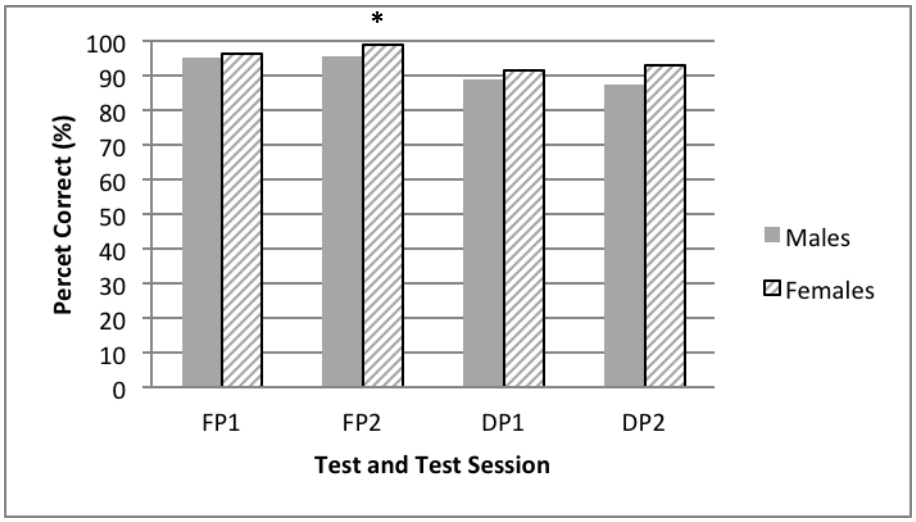

Figure 3. Mean scores for each test and test session for males denoted by solid fill and mean scores for each test and test session for females denoted by patterned fill. FP1 = Frequency pattern test session 1. FP2 = Frequency Pattern test session 2. DP1 = Duration Pattern test session 1. DP2 = Duration Pattern test session 2. ***** = significant difference between the mean scores of male and females ($p < 0.05$).

Figure 4. Mean score of males and females for the Gaps In Noise Test by test session

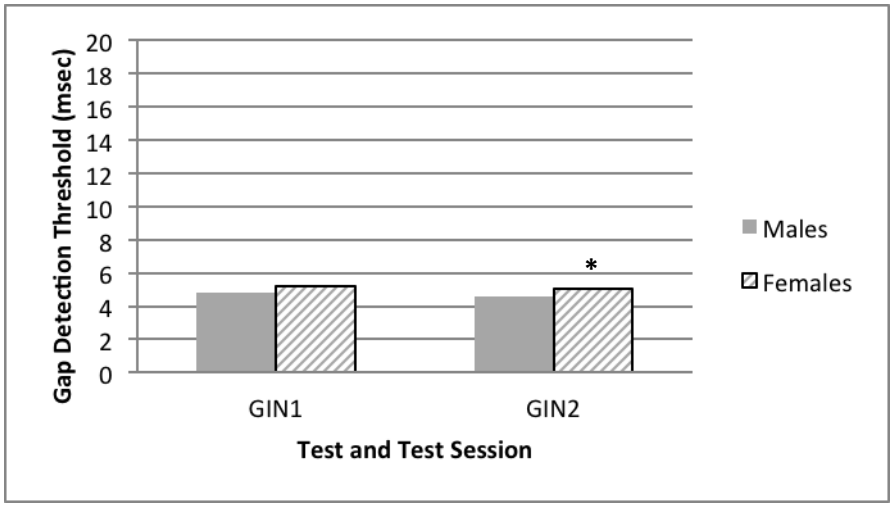

Figure 4. Mean scores for the Gaps In Noise test by test session for males denoted by a solid fill and females denoted by patterned fill. GIN1 = GIN T1. GIN2 = GIN T2. ***** = significant difference between the mean scores of male and females ($p < 0.05$).

Test-Retest Reliability.

A general linear model repeated measures ANOVA was completed individually for each test with gender as a between subject factor, test time as a within subject factor and scores as the dependent variable. Results indicated that there was no significance for test time, no significant effect of gender on scores, and no interaction between time and gender for any of the tests. The results from the general linear model repeated measures ANOVA can be seen in Table 8.

Table 8.

General Linear Model Repeated Measures ANOVA Outcomes for Overall Test Time, Effect of Gender on Scores and Interaction between Time and Gender for each Test.

	ANOVA Outcomes	
Test/Effect	F	Sig.
Frequency Pattern		
Time	3.52	.06
Gender	1.90	.17
Time x Gender	2.87	.09
Duration Pattern		
Time	.004	.95
Gender	2.54	.11
Time x Gender	1.64	.20
Gaps in Noise		
Time	1.27	.26
Gender	3.5	.65
Time x Gender	.14	.70

Note. F= F statistic. Sig.= p value.

In addition to analyzing the reliability of the FP, DP, and GIN test outcomes by evaluating group mean changes in scores, reliability was also analyzed by evaluating test-retest correlations for each test independently. The bivariate Pearson *r* correlations between test and retest scores for the FP, DP, and GIN test were r = .644, r = .710, and r = .449 respectively and were significant at the $p < 0.01$ level for all tests.

Clinical Significance

Although the aim of this study was not to quantify how many individuals obtained a "pass" or "fail" for each of the temporal processing tests, individual data was evaluated for clinical significance. As mentioned previously, clinically significant data would be any change in score that would a participant's score from a "passing" or "failing" score. For the purposes of this study and to evaluate clinical significance, the cutoff criteria for a passing score on the FP test was 80%, 72% on the DP test and a gap detection threshold of < 8 msec (Bellis, 2003; Musiek et al., 2005) For the FP, 32.6% (n =17) of ears had scores that improved on retest (T1<T2), 55.7% (n =29) had scores that maintained the same score (T1=T2), and 11.5% (n=6) had scores that were poorer on re-test (T1 > T2). Of all scores, 1.9% (n=1) of ears had a clinically significant change in scores that would have changed the clinician's diagnosis of APD. For the DP test, 28.5% (n=15) of ears had scores that improved on retest, 32.6% (n=17) had scores that maintained the same score, and 38.4% (n=20) had scores that were poorer on re-test. Of all the scores, 7.6% (n=4) of ears (3 participants) had a clinically significant change in score. Finally, for the GIN test, 26.9% (n=14) of ears improved on retest, 50.0% (n=26) maintained the same score, and 23.0% (n=12) had scores that were poorer on retest. Of note, 9.6% (n=5) of ears (5 participants) had a clinically significant change in score.

Individual data was also used to generate a scatterplot to illustrate the change in scores from T1 to T2 for each test for males and females and can be seen in Figures 5 and 6. It should be noted that scores were calculated by subtracting the score of T1 from the score of T2 (T2-T1), thus, if scores were better on re-test, this would result in a positive value; if scores were poorer on re-test, this would result in a negative value. For both

males and females, the change in scores ranged from -13.3 to 23.3% for FP, -30 to 16.6% for the DP, and -4 to 2 msec for the GIN test.

Figure 5. Change in male scores for the Frequency Pattern, Duration Pattern, and Gaps In Noise Test

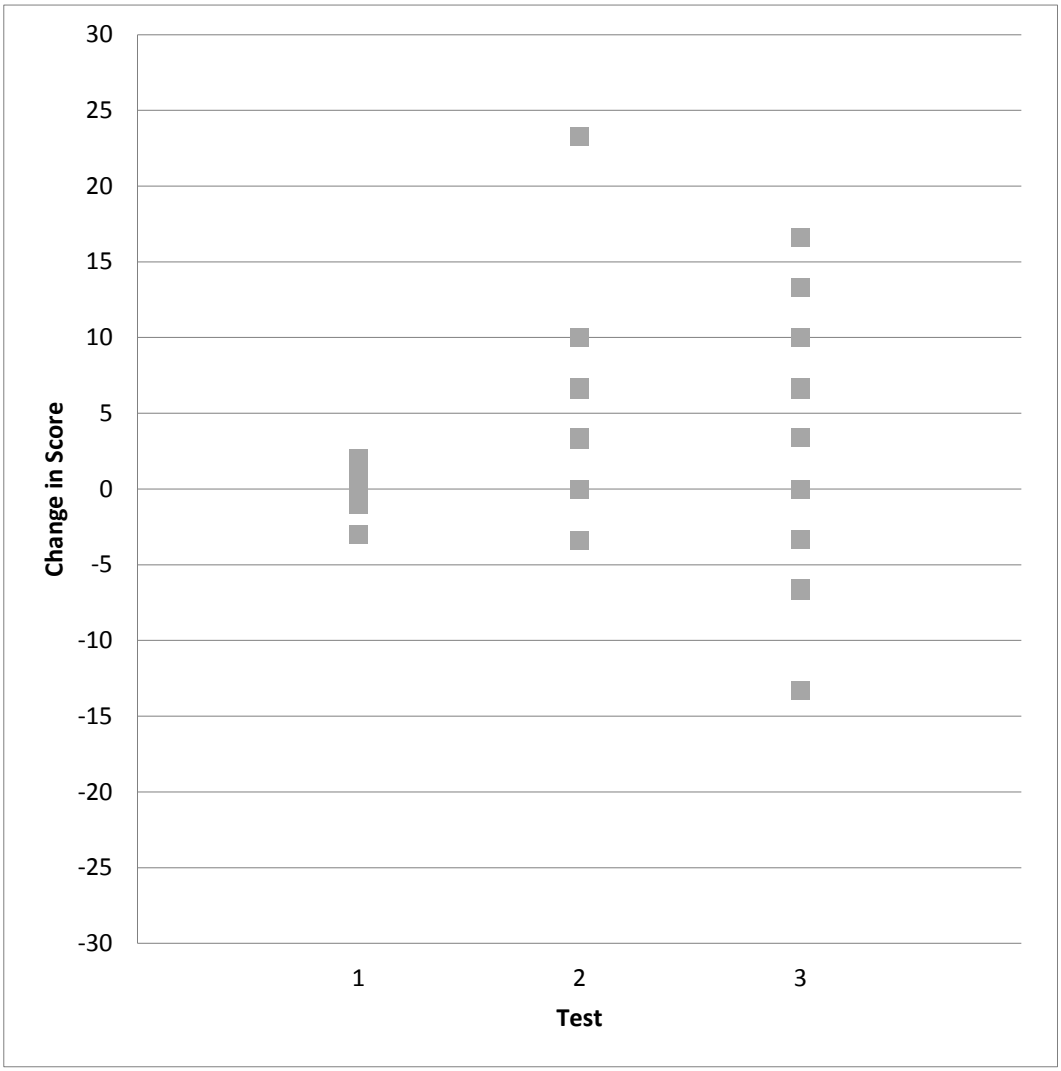

Figure 5. A scatterplot illustrating the change in score for male participants by calculating the difference between the scores of test session one from test session 2 (test session 2 – test session 1). Note: the change in scores for the Gaps In Noise test is a change in milliseconds while the change in scores for the Frequency Pattern and Duration Pattern test is the change in percentage. 1 = Gaps In Noise test. 2 = Frequency Pattern Test. 3 = Duration Pattern Test.

Figure 6. Change in female scores for the Frequency Pattern, Duration Pattern, and Gaps In Noise Test

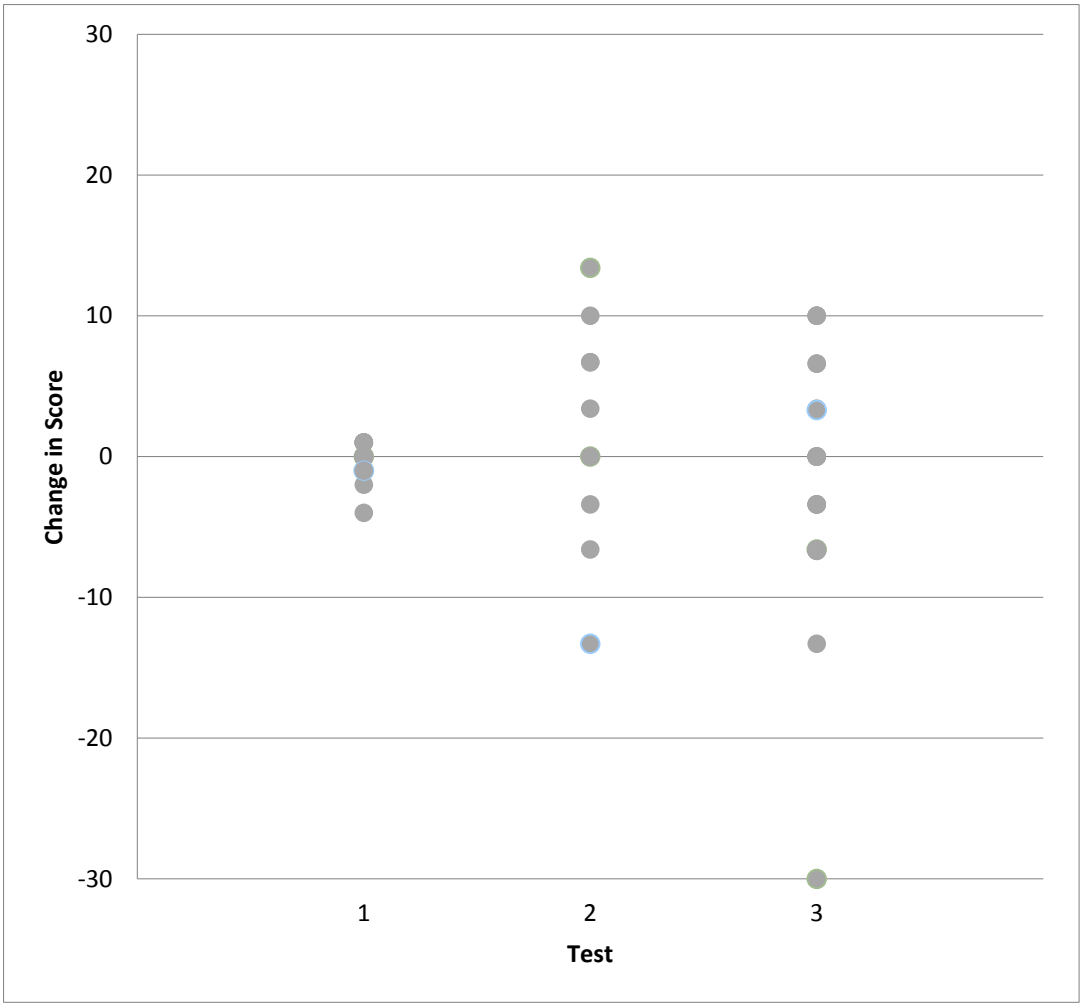

Figure 6. A scatterplot illustrating the change in score for female participants by calculating the difference between the scores of test session one from test session 2 (test session 2 – test session 1). Note: the change in scores for the Gaps In Noise test is a change in milliseconds while the change in scores for the Frequency Pattern and Duration Pattern test is the change in percentage. 1 = Gaps In Noise test. 2 = Frequency Pattern Test. 3 = Duration Pattern Test.

CHAPTER 5

Discussion

In the present study, temporal processing tests were evaluated to determine the test-retest reliability of three diagnostic tests (Frequency Pattern, Duration Pattern, and Gaps in Noise) in a group of 26 normal hearing adults (aged 20-29 years). These three temporal processing tests were administered to the participants at two different time periods referred to as time 1 (T1) and time 2 (T2). The time period between T1 and T2 was 7-9 days.

Temporal processing tests are a critical component of an APD test battery and should be utilized in the multidisciplinary evaluation of individuals with suspected APD, reading and/or language disorders, as well as with the elderly population in order to ensure accurate differential diagnosis and appropriate intervention therapies. Despite the functional importance of temporal processing abilities and the wide variety of populations that may be impacted by temporal processing deficits, to date, there is limited reliability data available for many of these commercially available tests (AAA, 2010; Strouse & Hall, 1995). Reliability data is critical for all tests within an APD test battery because APD is typically diagnosed within the clinical environment using behavioral measures which have inherent variability (Strouse & Hall, 1995). Test-retest reliability of APD measures has been historically difficult to assess in part due to normal changes that may occur in the participant's state during the central testing procedure. The participant's related factors that may influence their performance on an APD test include things such as their alertness during the test procedures and their motivation for completing the testing (Chermak & Musiek, 1997).

In the following section, the author will be presenting the test-retest reliability results for each of the three temporal processing tests and how they compare to the literature. Additionally, a discussion of the limitations of the present study and implications of these finding on clinical practice and future directions for research will be presented.

Test-Retest Reliability

Frequency Patterns. The results from the present study revealed that there was no significant difference between the mean scores on the FP test at T1 and T2. As expected, the mean FP test scores had a moderate and significant correlation ($r = 0.644$) across test sessions. No significant difference between the mean scores at T1 and T2 coupled with the fact that a moderate or better correlation was found suggests that the subjects' performance on the FP test was stable across the two test sessions and therefore, this test has moderate test-retest reliability.

The results related to the reliability of the FP test in the current study are in good agreement with results from previous studies within the literature (Humes, Coughlin, & Talley, 1996; Neijenhuis et al., 2011). Reliability results for the FP test as found by Neijenhuis et al. (2011) revealed that there was no significant difference between the mean FP test scores at T1 and T2 in normal hearing Dutch adults. The methodology of the Neijenhuis et al. (2011) study is important to note as the time between T1 and T2 ranged from 3.5-5.5 months. This is different from the methodology of the current study which utilized a time window of 7-9 days. The congruency of the reliability results for the FP test between the current study and the Neijenhuis et al. (2011) study despite

different re-test times is important because the retest time utilized in the Neijenhuis study is more typical of what occurs in the clinical environment. Thus, this indicates that even with extended time, performance on the FP test is stable.

Congruently, the reliability results for the FP test as reported by Humes et al. (1996) revealed that there was no significant difference between the mean FP test scores when using a re-test window of 15 minutes (time between T1 and T2) and 4 months (time between T1 and test session 3[T3]) in elderly patients who had normal to moderate sloping to high frequency SNHLs bilaterally. Additionally, Humes et al. (1996) reported a high correlation between the mean FP test score at T1 and T2 ($r = 0.92$) and moderate correlation at T1 and T3 ($r = 0.65$). The reliability results between T1 and T2 in the Humes et al. (1996) study are much better than the results found in the current study which was to be expected because of the short re-test time interval used; however, the reliability results found for T1 and T3 are in excellent agreement with the reliability results obtained in the current study despite several differences in the methodology which are important to note. The first major difference in the methodology of Humes et al. (1996) and the current study is the participant demographics. Humes et al. (1996) evaluated the reliability of the FP test in elderly patients (mean age of 73.3 years) who had varying degrees of hearing sensitivity. Despite documented hearing loss in all subjects, Humes et al. (1996) reported stable performance on the FP test, similar to the results obtained in the current study which evaluated young, normal hearing adults. This finding is important because it indicates that the FP test is a reliable clinical tool when it comes to assessing the temporal processing abilities of a wide range of populations, including the elderly population with varying degrees of hearing sensitivity. The second

major difference in the methodology between Humes et al. (1996) and the current study was the re-test time window. Humes et al. (1996) employed two different time windows, which were 15 minutes between T1 and T2 and 4 months between T1 and T3. Despite the extended duration of re-test time window utilized between T1 and T3, the reliability results were in good agreement with the reliability results obtained in the current study and again, indicate that performance on the FP test is stable regardless of extended re-test time windows which are more clinically relevant.

Collectively, the reliability results of the FP test obtained in the current study as well as those reported within previous studies, suggest that the FP test has moderate test-retest reliability and performance on this test is stable over an extended period of time (Humes et al., 1996; Neijenhuis, et al., 2011). These findings indicate that scores obtained at T1 accurately reflect the performance of the individual and there is little to no fluctuation in an individual's performance due to the test's design. Therefore, when utilizing the FP test within a clinical environment, any change in an individual's performance over time is likely due to true change in the individual's temporal processing skills.

Reliability data of other tests that are similar to the FP test have also been reported in the literature, including reliability results for different versions of the Pitch Pattern Test (PPT). The task of the PPT is identical to that of the FP test. However, there are differences in the test's design which include variations in stimulus frequencies, stimulus durations, and/or number of stimulus items. Two studies will be discussed that used different versions of the PPT and evaluated the pediatric population (Domitz & Schow, 2000; Summers, 2003).

Domitz and Schow (2000) evaluated the reliability of the Pitch Pattern 3 item test (Pinheiro, 1977) in seven third graders with a re-test window of 1-2 weeks. These researchers reported that a significant difference occurred in the mean scores of this version of the PPT at T1 and T2 and high correlation coefficients (r = 0.99). Thus, these reliability results indicate that performance on this version of the PPT is not stable in the pediatric population as Domitz and Schow (2000) reported that there was a learning effect. Similarly, Summers (2003) found comparable results as Domitz and Schow (2000) which evaluated the Pitch Pattern 4-item test in 19 third and fifth grade students. Summers (2003) reported a significant difference between the mean scores of T1 and T2 with a high correlation (r=.91) when a re-test window of 7-10 days was employed. Thus, the reliability results from Summers (2003) indicate that the 4-item version of the PPT is not stable over time in the pediatric population. The reliability results obtained by Domitz and Schow (2000) and Summers (2003) are not in good agreement with the reliability results of the FP test obtained in the current study despite the similar test-retest widow utilized across studies. The differences in reliability results between the Domitz and Schow (2000) and Summers (2003) studies and the current study may be due to the population tested (adults versus pediatrics), maturation and learning effects in the pediatric population, sample size, differences in the versions of the tests administered, and/or differences in the scoring of the tests as both Domitz and Schow (2000) and Summers (2003) scored reversals as correct (whereas reversals were considered incorrect in the current study).

Collectively, the reliability results reported in other studies employing other versions of the FP test indicate that these versions are not reliable tools to use clinically.

This is concluded due to the fact that there were significant differences between T1 and T2 which is an indicator of learning effects and therefore is not a good clinical tool. More research is needed on the reliability of the FP test in the pediatric population; however, the reliability results for the FP test utilized in the current study indicated that it has good test-retest reliability and has been shown to be a reliable tool in a diverse population of adults (Humes, et al., 1996; Neijenhuis, et al., 2011).

Duration Patterns. The results from the present study revealed that there was no significant difference between the mean scores on the DP test at T1 and T2. As expected, the mean DP test scores had a high and significant correlation ($r = 0.71$). This suggests that the subjects' performance on the DP test showed little variability across the two test sessions and therefore, the DP test has good test-retest reliability. Additionally, due to the lack of significance found between the mean scores at T1 and T1 and high correlation coefficients found, the DP test not only has good test-retest reliability, but it is also stable overtime.

The reliability results of the DP test in the current study are in good agreement with results from previous studies investigating DP test-retest reliability (Humes, et al., 1996; Neijenhuis, et al., 2011). Neijenhuis et al. (2011) evaluated test-retest reliability of the DP test and reported no significant difference between mean DP test scores at T1 and T2 in normal hearing Dutch adults. Since, the time window between T1 and T2 in the Neijenhuis et al (2011) study was 3.5-5.5 months, it appears that the DP test is not only reliable over a 7-9 day period, but also reliable over a longer time period, which is more typical of in the clinical environment. Thus, this indicates that performance on the DP test is stable over time and is a valuable tool within the clinic.

Similarly, Humes et al. (1996) reported no significant difference between mean DP test scores when using a re-test window of 15 minutes (time between T1 and T2) and 4 months (time between T1 and test session T3) in elderly patients who had normal to moderate sloping to high frequency SNHLs bilaterally. These researchers also reported a high correlation between the mean DP test score at T1 and T2 ($r = 0.97$) and at T1 and T3 ($r = 0.80$). These results are in good agreement with the results from the present study as high correlations were found across studies and suggests that the DP test is reliable over longer re-test periods than the ones used in the present study and therefore, useful in a typical clinical environment. Additionally, the results of the Humes et al. (1996) study suggest that the DP test is not only reliable when administered to young adults but also to elderly patients with varying degrees of hearing sensitivity (Humes et al., 1996).

Overall, the reliability results of the DP test in the current study and throughout the literature suggests that the DP test has good test-retest reliability and performance on this test is stable over time (Humes et al., 1996; Neijenhuis, et al., 2011). Due to the consistent report of high correlation coeffiecients and no differences in mean scores on the DP test across testing sessions indicates that scores obtained at T1 accurately reflect the performance of the individual and there is little to no fluctuation in an individual's performance due to the test's design. Due to this, any change in an individual's performance on the DP test overtime is likely due to true change in the individual's temporal processing skills and/or state of alertness.

Reliability data of another version of the DP test has also been reported within the literature. Summers (2003) evaluated the test-retest reliability of the 4 stimulus DP test in third and fifth grade students and utilized a test-retest time interval of 7-10 days. The

task of the 4 stimulus DP test is identical to that of the DP test discussed thus far, however, instead of presenting and asking the participant to repeat a pattern of three tones, the participants were presented and asked to repeat a pattern of four tones in order to control for ceiling effects (Summers, 2003). For this test, Summers (2003) reported no significant difference between the mean scores at T1 and T2 with a high correlation (r=.90). These reliability results are in good agreement with the reliability results for the DP test obtained in the current study. This is important because Summers (2003) recruited individuals from the pediatric population for testing and the results indicated that the performance on the 4-item version of the DP is stable over time in the pediatric population. Due to congruent reliability results obtained in the current study and those by Summers (2003), this indicates that the DP test is a reliable tool to utilize in the assessment of individuals of various ages.

Overall, the results reported within the literature and in the current study indicate that the DP test has good test-retest reliability (Humes et al., 1996; Neijenhuis, et al., 2011; Summers, 2003). These results also indicate that performance on the DP test is stable overtime in a wide range of populations including the pediatric, young adult, and elderly populations as well as in individuals with varying degrees of hearing sensitivity (Humes et al., 1996; Neijenhuis, et al., 2011; Summers, 2003). Thus, the DP test is a reliable measure of temporal processing abilities in a wide range of populations.

GIN. The results from the current study revealed that there was no significant difference between the mean scores on the GIN test at T1 and T2 when utilizing a re-test window of 7-9 days. The mean GIN test scores had a moderate and significant correlation (r = 0.449). This suggests that the GIN test has moderate reliability.

The reliability results of the GIN test in the current study are in essentially good agreement with results from a previous study within the literature (Musiek, et al., 2005). Musiek et al. (2005) evaluated the test-retest reliability of the GIN test and reported no significant difference between the mean GIN approximated gap detection threshold at T1 and T2 in normal hearing adults (aged 22-40 years) with a re-test window of 7-15 days. Despite a similar population pool and test-retest time interval, the Pearson product-moment correlations found by Musiek et al. (2005) were much better for left ($r = 0.95$) and right ($r = 0.88$) ears than those found in the current study. Although the correlation coefficient found in the current study for the GIN ($r = 0.449$) is much lower than the correlation values found by Musiek et al. (2005), the correlation coefficient obtained in the current study was still found to be significant at the $p < 0.05$ level. The differences in the results reported by Musiek et al. (2005) and in the current study may be associated with the differences in methodologies. Musiek et al. (2005) had a small sample size ($n=10$) of adults aged 22-40 years, whereas the current study had more than twice as many participants and a more restricted age range in order to control for temporal processing deficits which occur with age (Grose, et al., 2006; Humes, et al., 2010; Schneider, et al., 1994; Snell, 1997; Strouse, et al., 1998). More testing is needed in order to evaluate the reliability of the GIN as varied results have been obtained. However, due to the results of the current study and those reported by Musiek et al. (2205), at best, it can be concluded that the GIN has moderate test re-test reliability over time as no differences in mean scores was obtained and moderate to high correlations were found. Therefore, since the GIN had no significant difference between the mean scores and at least a moderate correlation coefficient, this test is stable over time.

Reliability and Other Audiologic Tests

The reliability results within the literature for the FP, DP and GIN tests will now be compared to the reliability results found for other APD tests that are used clinically. The FP, DP and GIN tests are tests of temporal processing and in general, studies within the literature have found that a majority of temporal processing tests have moderate to good test-retest reliability (Humes et al., 1996; Musiek et al., 2005; Summers, 2003). Temporal processing tests have been reported to have moderate to high correlation coefficients ($.65 < r < .97$) with no significant difference between the mean scores at T1 and T2 for tests such as the GIN (Musiek et al., 2005), DP test with three and four stimulus items (Humes et al., 1996; Summers, 2003), and the FP test (Humes et al., 1996). Thus, overall, temporal processing tests as a whole have been shown to have moderate to good reliability, which is in agreement with the reliability results found for the tests utilized in the current study.

Conversely, reliability results for dichotic tests have varied results. A few studies, which have investigated the test-retest reliability of dichotic listening tests, have indicated moderate to high correlations with no significant difference between the mean scores at T1 and T2 (Humes et al., 1996; Musiek, Gollegly, Kibbe, & Verkest-Lenz, 1991; Spencer, 2007; Strouse & Hall, 1995). Other studies have reported low to high correlations with significant differences between the mean scores at T1 and T2 (Amos & Humes, 1998; Domitz & Schow, 2000; Humes et al., 1996; Spencer, 2007; Summers, 2003). Some possible reasons for the discrepancy between the studies may be due to the dichotic tests utilized, the test-retest time window differences, the populations tested between the studies (pediatric versus adult), and/or maturation effects in the pediatric

population. Due to the confounding results found within the literature, it is hard to definitively state if dichotic tests have good test-retest reliability or not. Thus, as a whole, the reliability results reported in previous studies employing dichotic listening tests are not in agreement with the reliability results found in the current study or other related studies for temporal processing tests as a group. It is hypothesized that the differences in reliability results for the temporal processing tests and the dichotic tests may be due to the fact that these tests evaluate other areas of the brain that may be more susceptible to changes in maturation (especially in studies which evaluate the reliability of these tests in children), human alertness, differences in test design, and/or differences in study methodologies.

Overall, tests of monaural low redundancy have been shown in other studies to have poor test-retest reliability (Amos & Humes, 1998; Keith, 1986; Humes et al., 1996; Spencer, 2007; Summers, 2003). One study found that a few monaural low redundancy tests had high correlations ($.82 < r < .95$) with no significant difference between the mean scores at T1 and T2 (Humes et al., 1996). Other studies have found that many monaural low redundancy tests had poor correlations with no significant difference between the mean scores at T1 and T2 (Amos & Humes, 1998; Keith, 1986; Summers, 2003) or moderate to high correlations with significant differences between the mean scores at T1 and T2 (Amos & Humes, 1998; Keith, 1986; Humes et al., 1996; Spencer, 2007; Summers, 2003). Thus, as a whole, reliability results reported in other studies indicate that performance on monaural low redundancy tests is not stable over time and thus, these tests have poor test-retest reliability. These reliability results for monaural low redundancy tests are not in agreement with the reliability results found in the current

study or other studies investigating test-retest reliability of temporal processing tests. Again, it is hypothesized that the differences in reliability results for the temporal processing tests and monaural low redundancy tests may be due to the fact that some monaural low redundancy tests require more linguistic demands, evaluate a wide network of brain processes that may be more susceptible to changes in human alertness, differences in test design, and/or, differences in study methodologies.

An evaluation of the reliability results reported in other studies indicates that as a whole, temporal processing tests have the greatest test-retest reliability as compared to other categories of auditory processing tests. Performance on temporal processing tests has been reported to be the most stable over time as compared to the reliability results obtained in other studies for dichotic listening tests and monaural low redundancy tests. However, as a clinician, it is important to note how well reliability of temporal processing tests measure up to reliability of tests which are used in everyday clinical audiology centers to diagnose communication disorders.

Everyday Clinical Audiology Reliability. A brief review of pertinent studies indicates that pure tone testing as it is practiced currently has high reliability as pure tone thresholds have been reported to have no statistical difference between thresholds at T1 and T2 with a high correlation (r=.78) regardless of age, transducer type, and/or frequency tested (Schmuziger, Probst, & Smurzynski, 2004; Stuart, Stenstrom, Tompkins, & Vandenhoof, 1991). Additionally, immittance testing, including acoustic reflex testing and tympanometry, have been reported to have moderate to high correlations ($.69 < r < .95$) for all aspects of these tests when conducted in a manner typical of clinical practice (Margolis & Goycoolea, 1993; Mazlan, Kei, & Hickson, 2009;

Wiley & Barrett, 1991). Studies have shown that otoacoustic emission testing has moderate to high reliability for all frequencies tested with better correlation coefficients for the frequencies from 1-6 kHz for distortion product otoacoustic emissions ($.54 < r < .99$) and from 1-3 kHz for transient evoked otoacoustic emissions ($.78 < r < .92$) (Beattie, Kenwothy, & Luna, 2003; Franklin, McCoy, Martin, & Lonsbury-Martin, 1992; Wagner, Heppelmann, Vonthein, & Zenner, 2008). Thus, reliability data for the most common audiologic tests performed in the clinic on a daily basis indicates that these tests have good reliability. These reliability results are congruent with the reliability results found in the current study. This is important to note because although the reliability results from the current study do not reach the $r > .80$ cutoff criteria used in many audiology research studies, the reliability results of the temporal processing tests in the current study are congruent to the reliability results of everyday audiologic tests which are used to diagnose communication disorders (hearing loss).

Effects of Ear and Gender on Temporal Processing Abilities

Ear effect. Although not a primary aim of this study, effect of ear stimulated for each of the auditory processing tests was evaluated by test session. In this study, no significant difference between the mean scores for ear stimulated (left versus right) was found on any of the tests at either T1 or T2. These results were expected and agree with previous literature (DeFosse & Pinheiro, 1978; Musiek, et al., 1990; Musiek, et al., 2005; Musiek & Pinheiro, 1987; Zaidan, Garcia, Tedesco, & Baran, 2008).

Gender effect. A statistical evaluation of the means for male and female participants for each test session indicated a significant difference during T2 for the FP

and GIN tests. For the FP at T2, a small advantage in scores was recorded for females as their mean performance was 3.45% better than males. However, for T2 on the GIN test, males outperformed female participants by obtaining a mean gap detection threshold that was 0.5 msec better (lower). No other significant findings were obtained regarding male versus female scores.

The difference between males and females on the GIN test agrees with a study by Zaidan et al. (2008) which found that males performed significantly better than females with a mean gap detection threshold of 4.45 msec as compared to 5.61 msec. However, these authors note that the male participants were music therapy students and several studies have shown that temporal processing abilities are stronger in individuals with musical backgrounds (DeFosse & Pinheiro, 1978). Musical background of participants within this study is unknown.

The gender difference found for the FP test at T2 does not agree with data in the literature as many studies have reported no between-gender differences for the FP test (Jensen & Neff, 1993; Schochat & Musiek, 2006). The reason for the differences between the results of this study and the results reported within the literature are unknown but may be due to sample size differences of males and females between the studies. Additionally, another hypothesis as to why females outperformed males in this study on the FP test at T2 may be a result of a limitation of this study that will be discussed below regarding participant experience. In this study, all female participants may have had previous exposure to this test as all of these participants were either enrolled in the undergraduate Audiology, Speech-Language Pathology, and Deaf Studies department, the doctoral of Audiology department, or a faculty member of the audiology

department. Only five of 13 male participants may have experienced prior exposure to this test as they were enrolled in the doctor of audiology department. All other male participants were recruited from outside of the department(s) and likely had no previous exposure to the test.

Clinical Significance

Individual data was evaluated for clinical significance by evaluating the number of paired scores which had a change between test sessions that altered a "passing" or "failing" score on each test. This study utilized a cutoff criterion for a passing score on the FP test as an 80%, a 72% on the DP test and a gap detection threshold of < 8 msec in order to evaluate clinical significance (Bellis, 2003; Musiek et al., 2005). Of all scores, 1.9% (n=1), 7.6% (n =4), and 9.6% (n=5) of ears had a clinically significant change in scores on the FP, DP, and GIN test respectively. It should be noted that of all the ears that had a clinically significant change in scores on all of the tests (n = 10), 70% (n = 7 ears) showed improvement in scores which resulted in a change from a "failing" score to a "passing" score. One ear on the GIN and two ears on the DP test had a change in score which reflected a change from a "passing" score to a "failing" score. This indicates that when making a diagnosis of APD, interpretation of results from the FP, DP, and GIN tests should be made with caution as the results from this study show clinically significant changes in scores between test sessions for over a quarter of the participants in this study (n=7 participants). Additionally, since a majority of the ears which showed a clinically significant change in scores showed improved performance, interpretation of post-intervention assessments should be interpreted with caution as improvement in scores may be due to the inherent variability of human performance. However, it should

also be noted that a difference of scoring better by only one test item could result in a clinically significant change, as was the case with most of the participants in this study that had a clinically significant change in scores.

Study Limitations and Future Research

Sample Size. The aim of this study was to recruit 30 normal hearing participants between the ages of 18 and 35 years from the Baltimore, Maryland area. Because the primary aim of this study was to evaluate test-retest reliability data, testing took place in two different test sessions which were 7 to 9 days apart. This put the study at a greater risk for attrition as participants were required to return within a small window of time. As expected, data analyses could not be completed on 10% (n=3) of the participants as they failed to return for T2 and were lost to attrition. Additionally, one other participant was excluded due to a neurologic disorder. Thus, the final data analysis was completed on 26 participants. With a sample size this small, it is hard to generalize the findings from this study to the general population. Thus, future research should include a larger, more diverse sample size.

Participant experience. The testing for this study was completed at Towson University and the subject recruitment was aimed at young adults between the ages of 18 and 35. Therefore, a majority of the participants were undergraduate (n =7, 26.9%) and graduate (n=10, 38.4%) students enrolled in the Speech Language Pathology, Audiology, and Deaf Studies or Doctorate of Audiology program at Towson University. Additionally, one participant (n=1, 3.84%) was a professor of audiology at Towson University. Because of the course work in these programs, it is expected that these

participants (75% of all participants) had previous exposure to, or knowledge about, the tests that were administered. Thus, it is possible that the overall high performance on all of the tests could be contributed, at least in part, to previous exposure to the tests. If this is the case, scores obtained at T1 in this study may be higher than expected resulting in a reduced change in score between test time (reliability) than would be expected for more naïve listeners that make up the general population. Thus, it is hard to generalize these finding to the population at large and it is possible that these tests have greater variability (reduced reliability) between test sessions than those reported in this study. Future research in the area of test re-test reliability should include a more diverse sample including individuals with various educational and socioeconomic backgrounds.

Ceiling effects. As mentioned previously, the target participants for this study were normal hearing, normal developing young adults with no history or suspicion of APD, reading, language, and/or developmental disorders. Additionally, as discussed above, due to the location of the testing and the targeted age for participants, a majority of the participants (n=18, 75%) likely had previous exposure to the tests administered. Thus, although it was not a qualification of the study, essentially all participants performed within the normal, or "passing" range for both ears on all tests at T1 (FP n= 25, 96.1% DP n= 24, 92.3%, GIN n= 21, 80.7%). Additionally, over half (n=27, 51.9%) of the ears obtained a score of 100% on the FP at T1 and slightly less than half (n=21, 40.3%) of the ears only missed one test item on the DP test at T1. Thus, the scores on these tests did not have a normal distribution as they were skewed to the higher end of the distribution curve. Because of this, it was difficult to show any improvement in scores at T2 because the participants performed maximally at T1. It is difficult to relate ceiling

effects to the GIN due to the nature of the test; however, Phillips (1999) obtained minimum gap detection thresholds in highly trained normal hearing adults at 2 to 3 msec. However, other literature reports that minimum gap detection thresholds are much higher in less trained participants on the order of 5 msecs (Phillips & Smith, 2004). Given this information, over three quarters (n=45, 86.5%) of the ears in this study obtained a gap detection threshold of $\leq$ 5 msecs at T1. Therefore, future research should be conducted on individuals with a confirmed APD in order to have a more normal distribution of scores within a distribution curve. Retesting the participants can be conducted at the beginning of an aural rehabilitation appointment (to divert an ethical dilemma) or on participants that opt out of rehabilitation services and will enable to researchers to evaluate the true reliability of these tests without the potential confounds of ceiling effects.

Conclusions

An accurate diagnosis of APD is essential for rehabilitation purposes and it has been recommended and discussed within the literature that tests for APD should be sensitive, specific, and reliable. However, several studies have reported that this information is currently lacking within the literature (AAA, 2010; ASHA, 2005a; Jerger & Musiek, 2000; Strouse & Hall, 1995). APD is typically diagnosed based on behavioral tests and due to the inherent subjectivity of behavioral testing, test re-test reliability data is essential (Strouse & Hall, 1995). This data is essential because if tests used during the diagnosis of APD are reliable, then any observed change following intervention therapy may be attributed to an improvement in performance rather than to test re-test reliability errors (Theunissen et al., 2009; Wilson & McArdle, 2007). The purpose of this study

was to identify the test re-test reliability of three tests of temporal processing that have been found to be sensitive and specific (FP, DP, GIN) in normal hearing individuals in order to validate the use of these tests within an APD test battery. From a statistical standpoint, the results from this study indicated that the FP, DP, and GIN tests are reliable overtime. The statistical evaluation revealed that the DP test had the greatest test re-test reliability and the GIN had the poorest. From a clinical standpoint, when data was evaluated at the individual level, results indicated over a quarter of the participants (n=7 participants) had a clinically significant change in score that resulted in a change of "passing" or "failing" the tests between the test sessions. However, this information should be interpreted with caution as a change from "passing" to "failing" could result in performance of one test item better or worse on re-test. This is important to note and is another reason why administration of a battery of tests is essential when it comes to the diagnosis of APD and why a clinician should never diagnose APD based on the results of a single test. Overall, this study concludes that while the FP, DP, and GIN tests have shown to have moderate to high reliability results, the correlation results obtained were congruent with the reliability results of other audiologic tests which are used on a daily basis and are accepted as being valid diagnostic tools. Additionally, as a whole, temporal processing tests have the highest reliability results as compared to the peer-reviewed literature on the most common categories of behavioral APD tests. Therefore, this study concludes that the FP, DP, and GIN test are stable and reliable tests and have high clinical value.

APPENDIX

Appendix A: Institutional Review Board Approval

TOWSON UNIVERSITY

APPROVAL NUMBER: 11-A050

To: Stephanie Nagle
8000 York Road
Towson MD 21252

From: Institutional Review Board for the Protection of Human Subjects, Gerald Jerome, Member

Date: Wednesday, January 12, 2011

RE: Application for Approval of Research Involving the Use of Human Participants

APPROVED JAN 12 2011 TOWSON IRB

Office of University Research Services

Towson University
8000 York Road
Towson, MD 21252-0001

t. 410 704-2236
f. 410 704-4494

Thank you for submitting an Application for Approval of Research Involving the Use of Human Participants to the Institutional Review Board for the Protection of Human Participants (IRB) at Towson University. The IRB hereby approves your proposal titled:

Central Auditory Processing - Assessment and Rehabilitation

If you should encounter any new risks, reactions, or injuries while conducting your research, please notify the IRB. Should your research extend beyond one year in duration, or should there be substantive changes in your research protocol, you will need to submit another application for approval at that time.
We wish you every success in your research project. If you have any questions, please call me at (410) 704-2236.

CC:
File

TOWSON UNIVERSITY

Date: Wednesday, January 12, 2011

NOTICE OF APPROVAL

TO: Stephanie Nagle **DEPT:** AUDS

PROJECT TITLE: *Central Auditory Processing - Assessment and Rehabilitation*

SPONSORING AGENCY:

APPROVAL NUMBER: 11-A050

The Institutional Review Board for the Protection of Human Participants has approved the project described above. Approval was based on the descriptive material and procedures you submitted for review. Should any changes be made in your procedures, or if you should encounter any new risks, reactions, injuries, or deaths of persons as participants, you must notify the Board.

A consent form: [√] is [ ] is not required of each participant

Assent: [√] is [ ] is not required of each participant

This protocol was first approved on: 12-Jan-2011
This research will be reviewed every year from the date of first approval.

Gerald Jerome

Gerald Jerome, Member
Towson University Institutional Review Board

Appendix B: Consent Form for Participant in Research Project

Consent Form for Participation in a Research Project

Principal Investigator: Stephanie Nagle
Study Title: Central Auditory Assessment & Rehabilitation

1. Invitation to Participate
   You are invited to participate in a study of hearing by Dr. Stephanie Nagle of Towson University. Please read this form and ask any questions you may have before agreeing to be in the research study.

2. Purpose
   The purpose of this study is to help determine which tests among several are the best to diagnose certain types of hearing disorders.

3. Description of Procedures
   If you participate in this study, you will be required to listen to a variety of sounds such as tones, parts of words, words, and noises. You will be asked to tell us what you hear or press a button in response to what you hear. The study will include both normal and hearing impaired individuals. We need to test your hearing to understand whether the tests we perform during the study are valuable in diagnosing hearing problems. You may be excluded from the study if we find an ear infection or other types of conditions that may interfere with the tests. In some cases, small surface electrodes may be attached to your head or ear lobes with paste to record some responses to these various sounds. The electrodes are placed on top of the skin (i.e., ear, hand, neck, or clavicle) or scalp and do not hurt and paste is easily removed. For these kinds of tests, you will only have to sit quietly. The testing procedure may take from approximately one to two hours. Breaks from testing will be provided on a regular basis, and as requested. The experiments will take place at Towson University. An average experiment will take about 1.5 hours.

   You may be asked to fill out a questionnaire or case history related to hearing and communication difficulties. Your spouse or other communication partner may also be asked to fill out an auditory questionnaire about his/her perception of your hearing history and behavior.

4. Risks and Inconveniences

- WE believe there are no risks to you for your participation. No discomfort is associated with the task other than the usual fatigue or boredom related to sitting for an hour or two. The electrode cream used for some studies is non-toxic and washes off the skin easily, but if the electrodes are placed on the head, removal of all cream may require the use of shampoo at home.

5. Benefits

You may benefit directly from clinical assessment of your hearing, and central auditory processing abilities, and by therapeutic recommendations made based on those assessments. We hope this study may help to develop reliable tests and treatments which better diagnose and treat hearing disorders in the population as a whole.

6. Economic Considerations

You will not be paid nor will you be charged for participation in the study.

7. Confidentiality

The records from this study will be kept private. In any report published or presented regarding this study there will be no information that will reveal your identity. Records will be kept in a locked room and only researchers will have access to these records.

8. Voluntary Participation

You do not have to be in this study if you do not want to. If you agree to be in the study, but later change your mind, you may drop out at any time. There are no penalties or consequences of any kind if you decide that you does not want to participate. Your decision as to whether or not to participate will in no way affect any treatment at the Speech and Hearing Clinic or your student status at Towson University.

9. Do You Have Any Questions?

Take as long as you like before you make a decision. We will be happy to answer any question you or your child have about this study. If you have further questions about this project or if you child have a research-related problem, you may contact the principal investigator, Dr. Stephanie Nagle, at (410) 704-3554. If you have any questions concerning your rights as a research participant, you may contact Dr. Debi Gartland, Chairperson, Towson University Institutional Review Board (IRB), at 410-704-2236.

Authorization:
I have read this form and decided that I, _____ will
(name of subject)
participate in the project described above. Its general purposes, the particulars of involvement and possible hazards and inconveniences have been explained to my satisfaction.

Signature:_____

Date: _____

_____ _____
Signature of Primary Investigator Phone

THIS PROJECT HAS BEEN REVIEWED BY THE INSTITUTIONAL REVIEW BOARD FOR THE PROTECTION OF HUMAN PARTICIPANTS AT TOWSON UNIVERSITY.

Appendix C: Case History Form

1. Are you a native speaker of English/is English your primary language? Y / N
    i. (if not, stop here)
2. Have you ever been diagnosed with a hearing loss? Y / N
3. Do you suspect you have a hearing loss? Y / N
4. Have you ever had your hearing tested? Y / N
5. Do you have a history of chronic ear infections? Y / N
    i. If so, when was your last ear infection? _____
6. Have you ever been diagnosed with an auditory processing disorder? Y / N
7. Do you have a family history of hearing loss? Y / N
8. Do you have a history of excessive noise exposure (e.g. fire arms, profession in lawn services, military)? Y / N
9. Do you feel you have any difficulty with communication? Y / N
10. Have you had any surgeries to your head or neck? Y / N
    i. If so, when and for what? _____
11. Do you have a history/do you currently take an ototoxic medications (e.g. gentamyicin, loop diuretics, chemotherapy, hospitalized for an infection) Y / N
    i. If so, what are they? _____
12. Have you ever been diagnosed with a learning and/or language disorder? Y / N
    i. If so, what? _____
13. Do you suspect you have a learning and/or language disorder? Y / N
    i. Why? _____
14. Have you ever been diagnosed with any neurologic or degenerative disorder? Y / N
    i. If so, what? _____

References

Alho, K., Tervaniemi, M., Huotilainen, M., Lavikainen, J., Tiitinen, H., Ilmoniemi, R.J., & Naatanen, R. (1996). Processing of complex sounds in the human audiotry cortex as revealed by magnetic brain responses. *Psychophysiology, 33,* 369-375. Retrieved from http://becs.aalto.fi/mbl/pdf/alho_et_al_1996.pdf

American Academy of Audiology (AAA: 2010). Clinical practice guidelines: Diagnosis, treatment and management of children and adults with central auditory processing disorder. Retrieved from: http://www.audiology.org/resources/documentlibrary/Documents/CAPD%20Guidelines%208-2010.pdf

American Speech Language Hearing Association (ASHA: 2005). (Central) auditory processing disorders [Technical report]. Retrieved from http://www.asha.org/docs/html/tr2005-00043.html

Amos, N. E., & Humes, L. E. (1998). SCAN test-retest reliability for first- and third – grade children. *Journal of Speech, Language, and Hearing Research 41(4),* 834-845.

Bamiou, D-E., Musiek, F.E., & Luxon, L.M. (2001). Aetiology and clinical presentations of auditory processing disorders-a review. *Archives of Disease in Childhood, 85*(5), 361-365.

Baran, J.A., Bothfeld, R.W., & Musiek, F.E. (2004). Central auditory deficits associated with compromise of the primary auditory cortex. *Journal of the American Academy of Audiology, 15*(2), 106-116.

Beattie, R. C., Kenworthy, O. T., & Luna, C. A. (2003). Immediate and short-term reliability of distortion-product otoacoustic emissions. *International Journal of Audiology, 42*(6), 348-354.

Beck, D. L., & Bellis, T. J. (2010). *(Central) auditory processing disorders*: *Overview and amplification issues*. Oticon Library, Somerset: NJ

Bellis, T.J. (2003). *Assessment and management of central auditory processing disorders in the educational settings from science to practice* (2nd ed.). Clifton Park, NY: Thomson

Bellis, T. J. (2004, March). Redefining auditory processing disorder: An audiologist's perspective. *The AHSA leader.* Retrieved from http://www.asha.org/Publications/leader/2004/040330/f040330b1.htm

Bellis, T.J., & Ferre, J.M. (1999). Multidimensional approach to the differential diagnosis of central auditory processing disorders in children. *Journal of American Academy of Audiology 10*(6), 319–328.

Buonomano, D.V., & Karmarkar, U.R. (2002). How do we tell time? *The Neuroscientist, 8*(1), 42-51.

Cacace, A.T., & McFarland, D.J. (1998). Central auditory processing disorder in school-aged children: A critical review. *Journal of the Speech, Language, and Hearing Research, 41*(2), 355-373.

Carhart, R., & Jerger, J. J. (1959). Preferred method for clinical determination of pure-tone thresholds. *Journal of Speech and Hearing Disorders, 24*, 330-345.

Chedru, F., Bastard, V., & Efron, R. (1978). Auditory micropattern discrimination in brain damaged subjects. *Neuropsychologia, 16*(2), 141-149.

Chermak, G.D., & Lee, J. (2005). Comparison of children's performance on four tests of temporal resolution. *The Journal of the American Academy of Audiology, 16*(8), 554-563.

Chermak, G.D., & Musiek, F.E. (1997). *Central auditory processing disorders: New perspectives.* San Diego, CA: Singular.

Chisolm, T.H., Willott, J.F., & Lister, J.J. (2003). The aging auditory system: Anatomic and physiologic changes and implications for rehabilitation. *The International Journal of Audiology, 42*(Suppl 2), 2S3-2S10.

Colson, K., Robin, D., & Luschei, E. (1991). Auditory processing and sequential pitch and timing changes following frontal opercular damage. *Clinical Aphasiology, 20*, 317-325.

Cook, J.R., Mausbach, T., Burd, L., Gascon, G., Slotnick, H., Patterson, B., & Reynolds, B. (1993). A preliminary study of the relationship between central auditory processing disorder and attention deficit disorder. *Journal of Psychiatry and Neuroscience, 18*(3), 130-137.

Cox, L. C., McCoy, S. L., Tun, P. A., & Wingfield, A. (2008). Monotic auditory processing disorder tests in the older adult population. *Journal of the American Academy of Audiology, 19*(4), 293-308. doi: 10.3766/jaaa.19.4.3

Dawes, P., Bishop, D.V.M., Sirimanna, T., & Bamiou, D. (2008). Profile and aetiology of children diagnosed with auditory processing disorder (APD). *International Journal of Pediatric Otorhinolaryngology, 72*(4), 483-489. doi: 10.1016/j.ijporl.2007.12.007

DeFosse, C. O., & Pinheiro, M. (1978, November-December). *Perception of dichotic pitch patterns by musicians and non-musicians.* Paper presented at The Acoustical Society of America and The Acoustical Society of Japan Joint Meeting, Honolulu, HI. Abstract retrieved from http://dx.doi.org/10.1121/1.2003875

Domitz, D. M., & Schow, R. L. (2000). A new CAPD battery – multiple auditory processing assessment: Factor analysis and comparisons with SCAN. *American Journal of Audiology, 9*, 1059-0889. doi: 10.1044/1059-0889(2000/012)

Emanuel, D. C., Ficca, K. N., & Korczak, P. (2011). Survey of the diagnosis and management of auditory processing disorder. *American Journal of Audiology, 20,* 48-60. doi: 10.1044/1059-0889(2011/10-0019)

Ferre, J.M., & Wilber, L.A. (1986). Normal and learning disabled children's central auditory processing skills: An experimental test battery. *Ear and Hearing, 7*(5), 336-343.

Fitzgibbons, P.J., & Gordon-Salant, S. (1996). Auditory temporal processing in elderly listeners. *The Journal of the American Academy of Audiology, 7*(3), 183-189.

Franklin, D. J., McCoy, M. J., Martin, G. K., Lonsbury-Martin, B. L. (1992). Test/retest reliability of distortion-production and transiently evoked otoacoustic emissions. *Ear and Hearing, 13*(6), 417-429.

Frisina, R.D., & Walton, J.P. (2006). Age-related structural and functional changes in the cochlear nucleus. *Hearing Research, 217*, 216-223.

Gates, G.A., Anderson, M.L., Feeney, M.P., McCurry, S.M., & Larson, E.B. (2008). Central auditory dysfunction in older persons with memory impairment of Alzheimer dementia. *Archives of Otolaryngology, Head, and Neck Surgery,134*(7), 771-777.

Grimes, A.M., Grady, C.L., Foster, N.L., Sunderland, T., & Patronas, N.J. (1985). Central auditory function in Alzheimer's disease. *Neurology, 35*(3), 352-358.

Grose, J.H., Hall, J.W., & Buss, E. (2006). Temporal processing deficits in the presenescent auditory system. *The Journal of the Acoustical Society of America, 119*(4), 2305-2315.

Groth-Marnat, G. (2009). *Handbook of psychological assessment* (5th ed.). Hoboken, NJ: John Wiley & Sons, Inc.

Hall, J.W., & Johnston, K.N. (2007). Electroacoustic and electrophysiologic auditory measures in the assessment of (central) auditory processing disorder. In F.E. Musiek & G.D. Chermak (Eds.), Handbook of (central) auditory processing disorder: Auditory neuroscience and diagnosis (Vol. 1, pp. 287-317). San Diego, CA: Plural Publishing.

Helfer, K.S., & Vargo, M. (2009). Speech recognition and temporal processing in middle-aged women. *Journal of the American Academy of Audiology, 20*(4), 264-271. doi: 10.3766/jaaa..20.4.6

Hillyard, S. A., Hink, R. F., Schwent, V. L., & Picton, T. W. (1973). Electrical signs of selective attention in the human brain. *Science*, *182*(4108), 177-180.

Houtgast, T., & Festen, J.M. (2008). On the auditory and cognitive functions that may explain an individual's elevation of the speech reception thresholds in noise. *International Journal of Audiology, 47*(6), 287-295.

Huang, J., Yang, B., Zou, X., Jing, J., Pen, G., McAlonan, G.M., & Chan, R. (2012). Temporal processing impairment in children with attention-deficit-hyperactivity-disorder. *Research in Developmental Disabilities, 33*(2), 538-548.

Humes, L. E., Coughlin, M., & Talley, L. (1996). Evaluation of the use of a new compact disc for auditory perceptual assessment in the elderly. *Journal of the American Academy of Audiology, 7*, 419-427.

Humes, L.E., Kewley-Port, D., Fogerty, D., & Kinney, D. (2010). Measures of hearing threshold and temporal processing across the adult lifespan. *Hearing Research, 264*(1-2), 30-40. doi: 10.1016/j.heares.2009.09.010

Hurley, R., & Musiek, F. (1997). Effectiveness of three central auditory processing (CAP) tests in identifying cerebral lesions. *Journal of the American Academy of Audiology*, *8*(4), 257-262.

Jensen, J. K., & Neff, D. L. (1993). Development of basic auditory discrimination in preschool children. *Psychological Science, 4*(2), 104-107.

Jerger, J.F. (1970). Clinical experience with impedance audiometry. *Archives of Otolaryngology, 92*(4), 311-324.

Jerger, J., Jerger, S., Oliver, T., & Pirozzollo, F. (1989). Speech understanding in the elderly. *Ear and Hearing, 10*(2), 79-89.

Jerger, J., Johnson, K., Jerger, S., Coker, N., Pirozzolo, F., & Gray, L. (1991). Central auditory processing disorder: A case study. *Journal of the American Academy of Audiology, 2(1),* 36-54.

Jerger, J., & Musiek, F. (2000). Report of the consensus conference on the diagnosis of auditory processing disorder in school-aged children. *Journal of the American Academy of Audiology, 11*(9), 467-474.

Jirsa, R.E., & Clontz, K.B. (1990). Long latency auditory event-related potentials from children with auditory processing disorders. *Ear and Hearing, 11*(3), 222-232.

Kaga, K., Shindo, M., Tanaka, Y., & Haebara, H. (2000). Neuropathology of auditory agnosia following bilateral temporal lobe lesions: A case study. *Acta Otolaryngol, 120,* 259-262.

Karlsson, A., & Rosenhall, U. (1995). Clinical application of distorted speech audiometry. *Scandinavian Audiology, 24*(3), 155-160.

Keith, R. W. (1986). *SCAN; A Screening Test for Auditory Processing Disorders.* San Antonio, TX: The Psychological Corporation

Keith, R. W. (1995). Development and standardization of SCAN-A: Test of auditory processing disorders in adolescents and adults. *Journal of the Academy of Audiology, 6*(4), 286-292.

Keith, R. W. (2000). Development and standardization of SCAN-C test for auditory processing disorders in children. *Journal of the American Academy of Audiology, 11*(8), 438-445.

Keith, R. W. (2012a). Technical report: SCAN-3 for adolescents and adults: Tests for auditory processing disorders in adolescents and adults. Retrieved from http://pearsonassess.ca/hai/images/pdfs/scan-3-a-technical-report.pdf

Keith, R. W. (2012b). Technical report: SCAN-3 for children: Tests for auditory processing disorders in adolescents and adults. Retrieved from https://www.pearsonassessments.com/hai/images/pdf/technical_reports/6945A-2012_SCAN3C_TechReport.pdf

King, W.M., Lombardino, L.J., Crandell, C.C., & Leonard, C.M. (2003). Comorbid auditory processing disorder in developmental dyslexia. *Ear and Hearing, 24*(5), 448-456.

Kumar, A., & Sangamanatha, A.V. (2011). Temporal processing abilities across different age groups. *Journal of the American Academy of Audiology, 22*(1), 5-12. doi: 10.3766/jaaa.22.1.2

Lovett, B. J., & Johnson, T. L. (2010). Test review: R.W. Keith SCAN-3 for adolescents and adults: Tests for auditory processing disorders. San Antonio, Tx: Pearson, 2009. *Journal of Psychoeducational Assessment, 28*, 603-607. doi: 10.1177/0734282909353341

Lynn, G., & Gilroy, J. (1977). Evaluation of central auditory dysfunction in patients with neurological disorders. In: R. Keith , (Ed.), *Central auditory dysfunction* (pp. 177-221). New York, New York: Grune and Stratton.

Lynn, G.E., Gilroy, J., Taylor, P.C., & Leiser, R.P. (1981). Binaural masking-level differences in neurological disorders. *Archives of Otolaryngology, 107*(6), 357-362.

Margolis, R. H., & Goycoolea, H. G. (1993). Multifrequency tympanometry in normal adults. *Ear and Hearing, 14*(6), 408-413.

Martin, J.S., & Jerger, J.F. (2005). Some effects of aging on central auditory processing. *Journal of Rehabilitation Research and Development,42*(4), 25-44.

Mazlan, R., Kei, J., & Hickson, L. (2009). Test-retest reliability of the acoustic stapedial reflex test in healthy neonates. *Ear and Hearing, 30*(3), 295-301. doi: 10.1097/AUD.0b013e31819c3ea0

McArthur, G.M., & Bishop, D.V.M. (2004). Frequency discrimination deficits in people with specific language impairment: Reliability, validity, and linguistic correlates. *Journal of Speech, Language, and Hearing Research, 47*, 527-541. doi: 10.1044/1092-4388(2004/041)

Merzenich, M., Jenkins, W.M., Johnston, P., Schreiner, C., Miller, S.L., Tallal, P. (1996). Temporal processing deficits of language-learning impaired children ameliorated by training. *Science 271*(5245), 77-80.

Meyers, J.E., Roberts, R.J., Bayless, J.D., Volkert, K., & Evitts, P.E. (2002). Dichotic listening: expanded norms and clinical application. *Archives of Clinical Neuropsychology, 17*(1), 79-90.

Middelweerd, M.J., Festen, J.M. & Plomp, R. (1990). Difficulties with speech intelligibilities in noise in spite of a normal pure-tone audiogram. *Audiology, 29*(1), 1–7.

Mueller, H., Beck, W., & Sedge, R. (1987). Comparison of the efficiency of cortical level speech tests. *Seminars in Hearing, 8*(3), 279-298.

Musiek, F.E. (1983a). Assessment of central auditory dysfunction: The dichotic digits test revisited. *Ear and Hearing, 4*(2), 79-83.

Musiek, F.E. (1983b). Results of three dichotic speech tests on subjects with intracranial lesions. *Ear and Hearing, 4*(6). 318-323.

Musiek, F. E. (1994). Frequency (pitch) and duration pattern tests. *Journal of the American Academy of Audiology, 5*(4), 265-268.

Musiek, F.E. (2002) The frequency pattern test: A guide. *Hearing Journal, 55*(6), 58. doi: 10.1097/01.HJ.000029328/0.99394.dd

Musiek, F.E., & Baran, J.A. (2002). Central auditory evaluation of patients with neurological involvement. In J. Katz (Ed.), Handbook of clinical audiology (5th ed., pp. 532-544). Baltimore, MD: Lippincott Williams and Wilkins.

Musiek, F. E., & Baran, J. A. (2004). Audiological correlates to a rupture of a pontine arteriovenous malformation. *Journal of the American Academy of Audiology, 15*(2), 161-171.

Musiek, F.E., Baran, J.A., & Pinheiro, M.L. (1990). Duration pattern recognition in normal subjects and patients with cerebral and cochlear lesions. *Audiology, 29*(6), 304-313.

Musiek, F. E., Baran, J. A., Shinn, J.B., Guenette, L., Zaidan, E., & Weihing, J. (2007). Central deafness: An audiological case study. *International Journal of Audiology, 46*(8), 433-441.

Musiek, F., Charette, L., Kelly, T., Lee, W.W., & Musiek, E. (1999). Hit and false-positive rates for the middle latency response in patients with central nervous system involvement. *Journal of the American Academy of Audiology, 10*(3), 124-132.

Musiek, F. E., Charette, L., Morse, D., & Baran, J. A. (2004). Central deafness associated with a midbrain lesion. *Journal of the American Academy of Audiology, 15*(2), 133-151.

Musiek, F. E., & Chermak, G. D. (1994). Three commonly asked questions about central auditory processing disorders: Assessment. *American Journal of Audiology, 4,* 15-18.

Musiek, F.E., Chermak, G.D., Weihing, J., Zappulla, M., &Nagle, S. (2011). Diagnostic accuracy of established central auditory processing test batteries in patients with documented brain lesions. *Journal of the American Academy of Audiology, 22*(6), 342-358. doi: 10.3766/jaaa.22.6.4

Musiek, F.E., & Geurkink, N. (1982). Auditory brainstem responses and central auditory test findings for patients with brainstem lesions: A preliminary report. *Laryngoscope, 92*(8), 891-900.

Musiek, F. E., Gollegly, K. M., Kibbe, K. S., & Verkest-Lenz, S. B. (1991). Proposed screening test for central auditory disorders: Follow-up on the dichotic digits test. *American Journal of Otology*, *12*(2), 109-113.

Musiek, F.E., & Lee, W.W. (1995). The auditory brainstem response in patients with brain stem or cochlear pathology. *Ear and Hearing*, *16*(6), 631-636.

Musiek, F. E., & Pinheiro, M. L (1987). Frequency patterns in cochlear, brainstem, and cerebral lesions. *Audiology*, 26, 79-88.

Musiek, F., Pinheiro, M., & Wilson, D. (1980). Auditory pattern perception in split brain patients. *Archives of Otolaryngology, 106*(10), 610-612.

Musiek, F.E., Shinn, J., Hare, C. (2002). Plasticity, auditory training, and auditory processing disorders. *Seminars of Hearing, 23*(4), 263-272.

Musiek, F.E., Shinn, J.B., Jirsa, R., Bamiou, D.E., Baran, J.A., & Zaidan, E. (2005). The GIN (Gaps-in-Noise) Test performance in subjects with and without confirmed central auditory nervous system involvement. *Ear and Hearing, 26*(6). 608-618.

Neijenhuis, K.,Snik, A., & Broek, P. (2003). Auditory processing disorders in adults and children: evaluation of a test battery. *International Journal of Audiology, 42*(7), 391-400.

Neijenhuis, K., Stollman, M., Snik, A., & Van den Broek, P. (2001). Development of a central auditory test battery for adults. *Audiology, 40*(2), 69-77.

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

Noffsinger, D., Martinez, C. D., & Wilson, R. H. (1994). Dichotic listening to speech: Background and preliminary for digits, sentence, and nonsense syllables. *Journal of the American Academy of Audiology, 5*(4), 248-254.

Ostergard, C.A. (1983). Factors influencing validity and reliability of speech audiometry. *Seminars in Hearing, 4(3)*, 221-240.

Phillips, D. P. (1999). Auditory gap detection, perceptual channels and temporal resolution in speech perception. *Journal of the American Academy of Audiology, 10*(6), 343-354

Phillips, D.P. (2002). Central auditory system and central auditory processing disorders: Some conceptual issues. *Seminars in Hearing*, *23*(4), 251-261.

Phillips, D. P., & Smith, J. (2004). Correlations among within-channel and between-channel auditory gap detection thresholds in normal listeners. *Perception, 33*(3), 371-378.

Pinheiro, M. L. (1977). Test of central auditory function in children with learning disabilities. In R. Keith (Ed.), *Central auditory dysfunction* (pp. 223-256). New York, NY: Grune & Stratton.

Pinheiro, M.L., & Ptacek, P.H. (1971). Reversals in the perception of noise and tone patterns. *The Journal of the Acoustical Society of America*, *49*(6), 1778-1783.

Polite, D. F., & Beck, C. T. (2012). *Nursing Research:Generating and Assessing Evidence for Nursing Practice* (9th edition). Philadelphia: Lippincott Williams & Wilklins.

Rance, G., Corben, L., Barker, E., Carew, P., Chisari, D., Rogers, M … Delatycki, M. (2010). Auditory perception in individuals with Friedreich's Ataxia. *Audiology & Neurotology*, *15*(4), 229-240. doi: 10.1159/000255341

Rawool, V. W. (2007). Temporal processing in the auditory system. In D. Geffner &

Ross-Swain, D (Eds.). *Auditory Processing Disorders: Assessment, Management, and Treatment* (pp. 117-137). San Diego, CA: Plural Publishing.

Riccio, C.A., Hynd, G.W., Cohen, M.J., Hall, J., & Molt, L. (1994). Comorbidity of central auditory processing disorder and attention-deficit hyperactivity disorder. *Journal of the American Academy of Child & Adolescent Psychiatry, 33*(6), 849-857.

Samelli, A.G., & Schochat, E. (2008). The gaps-in-noise test: gap detection thresholds in normal-hearing young adults. *International Journal of Audiology, 47*(5), 238-245. doi: 10.1080/14992020801908244

Schmuziger, N., Probst, R., & Smurzynski, J. (2004). Test-retest reliability of pure-tone thresholds from 0.5 to 16 kHz using Sennheiser HAD 200 and Etymotic Research ER-2 earphones. *Ear and Hearing, 25*(2), 127-132.

Schneider, B.A., Pichora-Fuller, M.K., Kowalchuk, D., & Lamb, M. (1994). Gap detection and the precedence effect in young and old adults. *The Journal of the Acoustical Society of America, 95*(2), 980-991.

Schochat, E., & Musiek, F. E. (2006). Maturation of outcomes of behavioral and electrophysiologic tests of central auditory function. *Journal of Communication Disorders, 39*(1), 78-92.

Sharma, M., Purdy, S.C., & Kelly, A.S. (2009). Comorbidity of auditory processing, language, and reading disorders. *Journal of Speech, Language, and Hearing Research, 52*, 706-722.

Snell, K.B. (1997). Age related changes in temporal gap detection. *The Journal of the Acoustical Society of America, 101*(4), 2214-2220.

Spencer, M. L. (2007). *Reliability and test environment of the SCAN-A with children ages 12 – 15*. Unpublished manuscript, Department of Audiology, University of Maryland, College Park, Maryland. Retrieved from http://drum.lib.umd.edu/bitstream/1903/7263/1/umi-umd-4663.pdf

Stach, B.A., Loiselle, L.H., & Jerger, J.F. (1991). Special hearing aid considerations in elderly patients with auditory processing disorders. *Ear and Hearing, 12*, 131S-138S.

Strouse, A., Ashmead, D.H., Ohde, R.N., & Granthan, D.W. (1998). Temporal processing in the aging auditory system. *The Journal of the Acoustical Society of America, 104*(4), 2385-2399.

Strouse, A.L., & Hall, J.W. (1995). Test-retest reliability of a dichotic digits test for assessing central auditory function in Alzheimer's Disease. *Audiology, 34*(2), 85-90.

Strouse, A., Hall, J.W. 3rd, & Burger, M.C. (1995). Central auditory processing in Alzheimer's disease. *Ear & Hearing, 16*(2), 230-238.

Stuart, A., Stenstrom, R., Tompkins, C., & Vandenhoff, S. (1991). Test-retest variability in audiometric threshold with supraaural and insert earphone among children and adults. *Audiology, 30*(2), 82-90.

Summers, S. A. (2003). *Factor structure, correlations, and mean data on form A of the BETA III version of multiple auditory processing assessment (MAPA)*. Unpublished

manuscript, Department of Communication Sciences and Education of the Deaf, Idaho State University, Pocatello, Idaho.

Tallal, P. (1980). Auditory temporal perception, phonic, and reading disabilities in children. *Brain Language, 9*(2), 182-198.

Tallal, P., & Piercy, L. (1973). Defects of non-verbal auditory perception in children with developmental aphasia. *Nature, 241,* 468-469.

Theunissen, M., Swanepoel, D.W., & Hanekom, J. (2009). Sentence recognition in noise: Variables in compilation and interpretation of tests. *International Journal of Audiology, 48*(11), 743-757. doi: 10.3109/14992020903082088

Tremblay, K., Kraus, N., McGee, T., Ponton, C., & Otis, B. (2001). Central Auditory Plasticity: Changes in the complex after speech-sound training. *Ear and Hearing, 22*, 79-90.

Turgeon, C., Champoux, F., Lepore, F., Leclerc, S., & Ellemberg, D. (2011). Auditory processing after sport-related concussions. *Ear and Hearing, 32*(5), 667-670.

Wagner, W., Heppelmann, G., Vonthein, R., & Zenner, H.P. (2008). Test-retest repeatability of distortion product otoacoustic emissions. *Ear and Hearing, 29*(3), 378-391. doi: 10.1097/AUD.0b013e31816906e7

Walker, M., Shinn, J., Cranford, J., Givens, G., & Holbert, D. (2002). Auditory temporal processing performance of young adults with reading disorders. *Journal of Speech, Language, and Hearing Research. 45,* 598-605.

Wiley, T. L., & Barrett, K. A. (1991). Test-retest reliability in tympanometry. *Journal of Speech and Hearing Research, 34*(5), 1197-1206.

Wilson, R.H., & McArdle, R. (2007). Intra- and inter-session test, retest reliability of Words-in-Noise (WIN) test. *Journal of the American Academy of Audiology, 18*(10), 813-825.

Witton, C., (2010). Childhood auditory processing disorder as a developmental disorder: The case for a multi-professional approach to diagnosis and management. *International Journal of Audiology, 42*(2), 83-87.

Wright, B.A., Buonomano, D.V., Mahncke, H.W., & Merzenich, M.M. (1997). Learning and generalization of auditory temporal-interval discrimination in humans. *Journal of Neuroscience, 17*(10), 3956-3963.

Zaidan, E., Garcia, A. P., Tedesco, M. L., & Baran, J. A. (2008). Performance of normal young adults in two temporal resolution tests. *Pró-Fono Revista de Atualização Científica*, *20*(1), 19-24.

CURRICULUM VITA

NAME: Crystal Lilly, B.S.

PERMANENT ADDRESS: 1781 Long Corner Road

Mount Airy, MD 21771

PROGRAM OF STUDY: Doctorate of Audiology

DEGREE AND DATE TO BE CONFERRED: Doctorate of Audiology, 2014

Secondary education: Glenelg High School, Glenelg, MD, 2006

Collegiate institutions attended	Dates	Degree	Date of Degree
Towson University	2006-2010	Bachelor of Science	2010
Towson University	2010-Present	Doctorate of Audiology	2014

Major: Doctorate of Audiology

Audiology, Speech Language Pathology, and Deaf Studies

Minor: Psychology

Professional positions held:

ENTTA Care
Chesapeake Hearing Centers
Baltimore County School Systems
Chesapeake Ear, Nose, & Throat, P.A.