Spring 2017

A Vowel Analysis of the Northwestern University-Children's Perception of Speech Evaluation Tool

Kassie Nicole Zukowski
University of New Hampshire, Durham, knr37@wildcats.unh.edu

A VOWEL ANALYSIS OF THE NORTHWESTERN UNIVERSITY – CHILDREN'S PERCEPTION OF SPEECH EVALUATION TOOL

by

KASSIE ZUKOWSKI

Honors Thesis

Submitted to the University of New Hampshire in Partial Fulfillment of the Requirements for the

Honors in Major for the Degree of Baccalaureate of Science in

Communication Sciences and Disorders

May 2017

Abstract

In an analysis of the speech perception evaluation tool, the Northwestern University – Children's Perception of Speech test, the goal was to determine whether the foil words and the target word were phonemically balanced across each page of test Book A, as it corresponds to the target words presented in Test Form 1 and Test Form 2 independently. Based on vowel sounds alone, variation exists in the vowels that appear on a test page on the majority of pages. The corresponding formant frequencies, at all three resonance levels for both the average adult male speaker and the average adult female speaker, revealed that the target word could be easily distinguished from the foil words on the premise of percent differences calculated between the formants of the target vowel and the foil vowels. For the population of children with hearing impairments, especially those with limited or no access to the high frequencies, the NU-CHIPS evaluation tool may not be the best indicator of the child's speech perception ability due to significant vowel variations.

Keywords: hearing impairment, vowel, variation, formant frequency

Introduction

Pediatric hearing loss is a disorder in which a child or adolescent is unable to detect or distinguish the range of sounds naturally available to the human ear (Justice & Redle, 2014). According to the National Institute on Deafness and Other Communication Disorders (2016), almost three of every 1,000 children are born with a detectable level of hearing loss in one or both ears, with over half resulting from genetic causes, in the United States. In addition to hereditary sources, pediatric hearing loss can be developmental, stemming from middle ear pathologies, such as otitis media, or an ear infection. Approximately five out of six children experience ear infections by the time they are three years old and often times multiple occurrences leave lasting damage on the child's auditory system (National Institute on Deafness and Other Communication Disorders, 2016).

Although relatively few school-aged children, roughly one to two percent, exhibit severe or profound permanent hearing loss, a much larger number of students portray hearing loss that is serious enough to impact their educational achievement. Nearly 20 percent of school children have a hearing loss that is considered significant enough to impede the child's academic performance (Shargordosky, Curhan, Curhan, & Eavey, 2010; Justice & Redle, 2014). Whether the hearing loss is present at birth or develops sometime thereafter, it can have a substantial influence on the language acquisition and language ability of the child (Martin & Clark, 2009). For this reason, audiologists, speech-language pathologists, and educators each play an important role in diagnosing and treating children with hearing loss (Justice & Redel, 2014). Among the many avenues to explore when hearing loss is suspected is the child's speech perception abilities, as it provides evidence of what the child can hear and raises questions about what the child is unable to hear.

I. Speech Perception

While hearing accounts for acknowledging the presence of sound, it is what a listener does next that provides researchers and clinicians with critical information to truly understand the way which the auditory system functions. How a listener perceives the sound wave that stimulates the auditory nerve assigns meaning to the utterance that was heard. The meaning of speech is derived from the ability to discriminate between and identify the acoustic-phonetic aspects of the incoming word or sentence. Although most listeners are unaware of the constant, ongoing process, the perception of speech depends on the listener's ability to segment a continuously changing stream of sound into its parts, based on the variation in meaning (Ferrand, 2014).

The numerous sounds in the English language can be combined in countless ways to form the building blocks of connected speech. From syllables to words to phrases, listeners extract meaningful information from what is heard. Despite this segmenting process that occurs through active cognitive processing, the sound waves always remain consistent; in other words, a single phoneme does not carry the same linguistic information from one context to another (Tatham & Morton, 2011; Ferrand, 2014). Speech perception becomes even more complicated because phonemes can vary immensely in their production for a number of reasons including differences in production from speaker to speaker, and the fact that phonemes are influenced by their surrounding phonemes; preceding and following sounds affect phonemes, creating coarticulation, which plays a role in the perception of speech (Ferrand, 2014).

Since phoneme production can fluctuate with rate of speech, patterns of stress, changes in intonation, and dialect, an individual phoneme cannot be linked to a single acoustic responsibility (Ferrand, 2014). This explains why the study of speech perception has caused challenges in the realm of research for academics as well as practicing clinicians in the fields of speech-language pathology and audiology (Madell, 1998). Much like the components of speech that complicate the concept of speech production on the account of the speaker, other factors exist that play a role in how the listener can perceive spoken word; these considerations include the listener's ability to recognize the phonemes, syllables, and prosodic patterns of his or her language, knowledge of the phonological, syntactic, semantic, and pragmatic features of his or her language, and ability to detect and interpret the emotional content of a message (Tatham & Morton, 2011). These features not only play a role in how the information inherits meaning by the listener, but it also influences how evaluation of speech perception functions.

II. Speech Perception in Children

Speech perception data provides a great deal of information related to assessing a child's higher level auditory processing abilities as well as evaluating intervention outcomes. In the pediatric population, speech perception measures assist in determining whether a child is benefiting from his or her hearing device, whether it be a hearing aid or a cochlear implant. This information can also provide insight into whether a child with hearing aids should be considered for a cochlear implant. Additionally, speech perception examinations help track performance over time as it provides a point of reference and can be replicated to determine improvements or setbacks. Finally, when combined with speech and language outcomes, speech perception data is critical for establishing guidelines for habilitation (Eisenberg, Johnson & Martinez, 2005). Since it has the ability to provide a wide range of information, spanning a relatively long period of time, speech perception is often used in the evaluation of children with hearing loss. Despite being a critical component in understanding auditory perceptual ability, the field is lacking a single, reliable evaluation tool that can be widely supported and thus, a variety of tools exist, leaving the approach to the discretion of clinicians.

III. Speech Perception Assessments

Over the last several decades, speech perception has gained significant attention from those in the fields of speech-language pathology and audiology as it measures more than a listener's ability to hear; it helps to gain insight as to how the listener assigns meaning based on combining the sense of hearing with a comprehension task. Evaluations utilizing speech stimuli can assist in the discovery of what kind of impact the hearing impairment has on the listener's ability to recognize and discriminate speech sounds. This information can then determine the type and severity of the hearing disorder, provide information regarding whether an assistive sensory aid is necessary and if so, what form it should take (either a hearing aid or cochlear implant), inform clinicians of rehabilitation strategies, and estimate a child's listening abilities in real-word situations (Dobie & Hemel, 2004).

In pediatric clients, Cienkowski, Ross, and Lerman (2009) brought attention to three characteristics that are essential for an accurate evaluation of a child's speech perception abilities; the test must include speech material that falls within the child's receptive vocabulary, the mode of response must be age-appropriate, and reinforcement must be readily available if necessary. Even with the utmost care, it is likely that the results of an evaluation may be influenced by the examinee's attention, interest, and motivation (Norther & Downs, 2002). Since a single standardized test for speech perception has yet to be determined, many options exist (Cienkowski, Ross & Lerman, 2009). These examinations take a number of forms, cater to a variety of age groups, and obtain a wide-range of information.

i. Closed Tests

Of the many testing styles, closed-set tests are often used in the examination of auditory perceptual ability. Restricting the participant to a single response out of a fixed number of possible options, similar in style to multiple choice testing, closed-set tests reduce the cognitive demands that can often arise in speech perception tasks (Dobie & Hemel, 2004). Especially in early childhood, when language is just beginning to develop, utilizing a limited vocabulary is a method that can allow the evaluation to be extended across varying ability levels. This format can be especially fitting in determining speech recognition in listeners with reduced language skills or limited speaking and writing abilities. Even though it meets the needs of a wide range of children, a closed-set format has the potential to inadequately evaluate the perceptual processes that are performed in daily life (Dobie & Hemel, 2004). The delivery of these evaluations are often not representative of the natural speech production and processing that occurs in natural conversations in real-life situations.

When exploring the area of perception in children, evaluation is often done at the word level to match the child's ability to use and understand language. Although it is not necessarily representative of everyday conversation, it provides a baseline of information that can help researchers and clinicians to pinpoint how successful a child is at assigning meaning to what is heard. Based on the accuracy of the results, professionals obtain information regarding perception and hearing level (Alpiner & McCarthy, 2000). While single words are at the premise of the exam, information can be obtained regarding particular speech sounds. Just as listeners segment connected speech into individual phonemes, clinicians can follow a similar process to discover what the child can and cannot hear.

Word recognition tests evaluate a child's ability to understand speech under different listening conditions. Unlike threshold testing, word recognition testing is performed at suprathreshold levels. Evaluation may be conducted at different intensities and under various conditions of competing noise to extend the data obtained beyond speech perception abilities. The selection of test materials and conditions will depend on the child's vocabulary level and his or her ability to cooperate. The scoring element demands the use of two categories, accurate perception and inaccurate perception of all the phonemes in the target word. By modifying the response task and types of reinforcement, it is possible to learn a great deal about a child's speech perception skills, without which information an audiological evaluation is incomplete (Madell, 1998).

IV. Vowels

Vowels are phonemes that are produced without any substantial constriction of air flow in the vocal tract. Even though only five Roman alphabet letters exist to orthographically represent the vowels of the English language, many more vowel sounds can be produced in spoken language. Without any blockage of the oral cavity, the tongue acts as the primary articulator. With its muscular attachments to the mandible, the tongue functions in partnership with another oral component, the jaw, in the production of vowel sounds. While these structures actively move in position to produce the desired sound, the byproduct is an alteration of the shape and size of the pharynx (Small, 2012). With the addition of air flow, vowels become audible. While these highly specific actions of the tongue and mandible are the main articulators used to produce vowels, they are also directly linked to the categorization system that vowels follow (Small, 2012).

In the traditional classification system, vowels are classified according to two observable characteristics of the vowels' primary articulator, the tongue. Unlike the organizational system of consonants, all vowels are voiced, eliminating the need for a voicing label. Since all vowels are produced with a relatively open tract, the manner category does not apply to vowels. However, the place of articulation does vary from vowel to vowel, which is seen by the height of the tongue or how high or low the body is held within the oral cavity. Additionally, tongue advancement, when discussing how advanced or how retracted the tongue body is within the oral cavity, also changes as different vowels are produced (Ferrand, 2014). These dimensions, tongue height and tongue advancement, are the most significant factors to impact vowel production, but other characteristics play a role, such as how tense or relaxed the tongue is as well as the amount of rounding that occurs at the lips. Although minor in comparison to tongue position, these features contribute to accurate production of vowels (Raphael, Borden, & Harris, 2007).

Vowels are often represented schematically in what is commonly known as the vowel quadrilateral. Taking the form stated in its name, the shape represents the oral cavity; the vertical position of each vowel sound corresponds with the highest point of the tongue during articulation of the vowel while the horizontal position signifies the degree of tongue advancement within the oral cavity. Additional subcategories are placed in the vowel quadrilateral to provide immediate information about the vowel sounds. In regard to tongue height, vowels can be *high, mid,* or *low.* Tongue advancement can be divided further into *front, central,* or *back* (Ferrand, 2014). It is this system and visual representation that create the foundation of vowel classification, providing a simple way for these speech sounds to be described and compared.

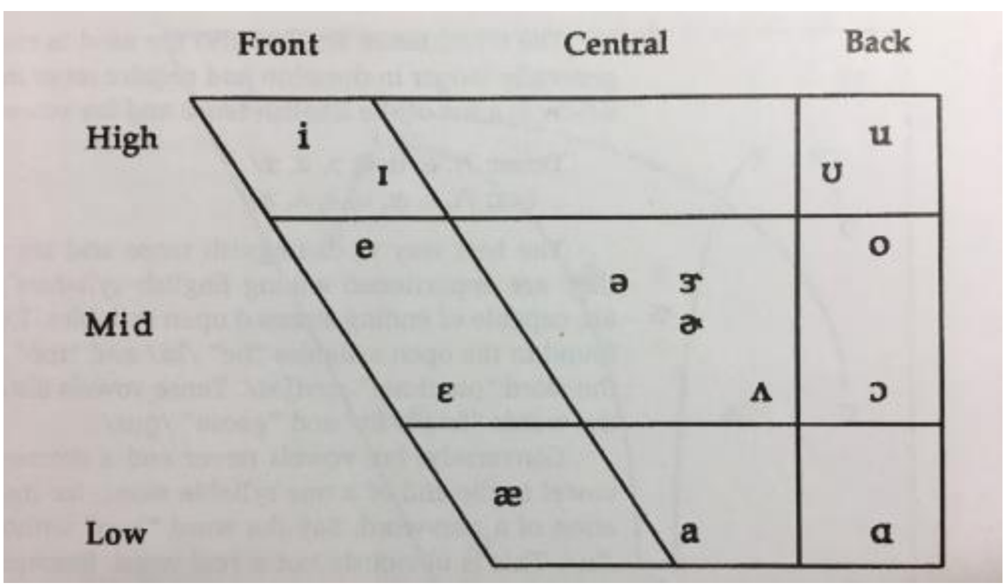

Figure 1: Vowel Quadrilateral (Small, 2014)

i. Acoustic Characteristics of Vowels

The four corner vowels, those located at each edge of the quadrilateral, (/i/, /u/, /a/, /æ/), are produced with the tongue as far from the neutral tongue position as possible, therefore representing the limit of vowel articulation. These periodic complex sounds are characterized by their first three formants (Ferrand, 2014). A formant is defined as the peak of energy in the spectrum of a vowel sound, caused by the narrowing and widening of the vocal tract and resulting in frequency and intensity changes. The peaks are numbered consecutively, beginning with the lowest formant, formant one (F1), formant two (F2), and the highest formant, formant three (F3). It is these formants and the manipulation of formant frequencies that prompt the recognition of different vowel sounds in listeners (Martin & Clark, 2009).

Monophthongs, the vowels depicted in the vowel quadrilateral, are produced with a relatively constant tongue position. Both front and back vowels have common characteristic formant patterns. For front vowels, F1 and F2 are spaced very widely apart whereas F2 and F3 are much closer together. Ferrand (2014) shares the example of the highest front vowel /i/ that follows this general rule; the greatest degree of separation is between F1 and F2 and all consecutive, lower front vowels decrease in separation between these two formants. The back vowels, however, are classified differently; F1 and F2 are closely spaced, with F3 reading a much higher frequency. As for central vowels, these frequencies tend to show a more even balance, with an equally spaced relationship between all three formants.

Not all vowel sounds are pure in nature, maintaining one tongue position throughout. These speech sounds are known as diphthongs and are produced by uttering two vowel sounds as one unit. Unlike monophthongs, diphthongs are vowels in which their resonance characteristics change during production. The onglide and offglide, the starting point of the first vowel and the ending point of the second vowel respectively, are not the same, causing the formants to shift in frequency from the beginning of the sound to the end of the sound. These inconsistencies are what is known as formant transitions (Ferrand, 2014). Since these double vowel sounds are linked so closely together, not only do they show an adjustment in frequency from beginning to end, but they also display a longer duration than that of their monophthong partners (Small, 2012).

For example, the diphthongal allophone of /e/ is /eɪ/, which is made up of the onglide /e/ and ends with the offglide /ɪ/. Commonly used as the phonetic representation that is equivalent to /e/ in stressed syllables or when the vowel sound occurs at the end of the word, regardless of stress, the diphthong /eɪ/ is articulated with the tongue in the initial position for the vowel /e/ then it glides to the high, front position for /ɪ/. Although seemingly a subtle adjustment to tongue position in terms of production, the formant transition varies significantly, increasing by approximately 1000 Hz from beginning to end (Small, 2012).

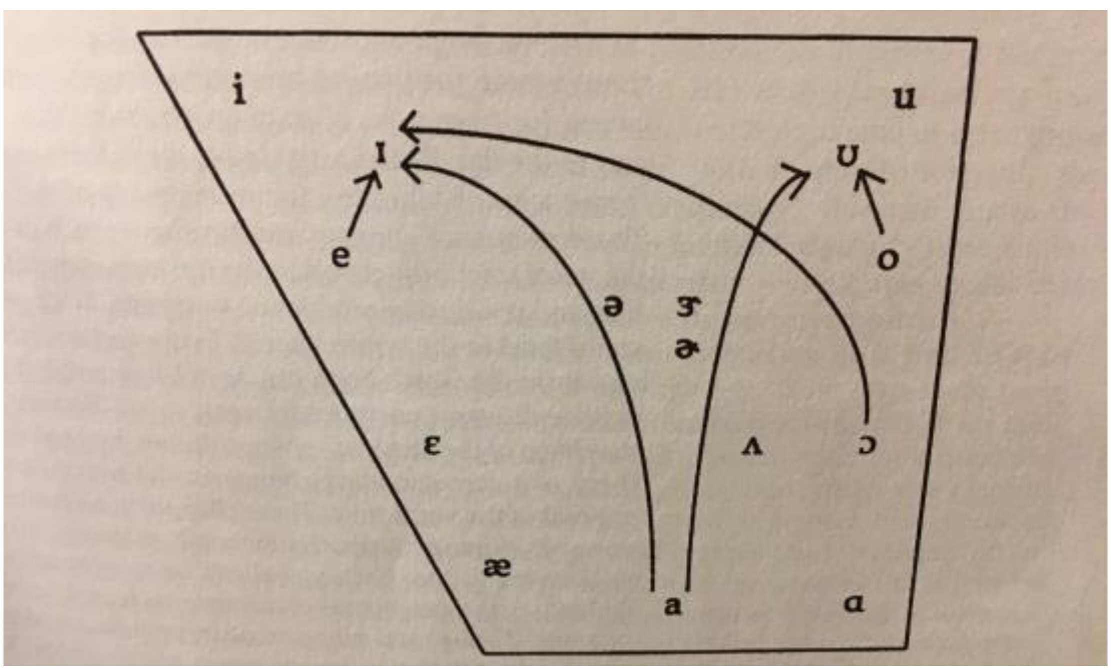

Figure 2: Onglide and Offglide Position for Diphthongs (Small, 2012)

V. Perception of Vowels

From their categorization, it is apparent that vowels rely entirely on the shape and size of the vocal tract for proper production. It is the position of the articulators within the oral cavity and the byproduct of resonance that provides each vowel with its own distinct frequencies. The three level formant system and how these formants relate to one another are the two determining factors impacting vowel distinction by the listener (Ferrand, 2014).

In comparison to consonants, vowels are both longer in duration and more intense. Their frequency components are typically lower than the frequency components of consonants. The three formant frequencies that are critical to the production and perception of vowels do not exist in the same way for consonants, making these nucleus sounds more challenging to work with. In research conducted by Strange (1999), each formant does not play the same role when it comes to how a listener perceives the sound. That is, all three frequencies are not equal in speech perception. As discussed earlier, the F1 and F2 values for back values are much closer than that of F3 whereas front vowels function in an opposite manner, where F1 is distinctly distant from F2 and F3, which are relatively close together. When these formant frequencies are close together, such as F1 and F2 for back vowels and F2 and F3 for front vowels, the listener integrates the two elements into one perceptual unit, equivalent to an average of the two closely spaced formant frequencies. Thus, back vowels are perceived on the basis of an F1 and F2 average and its F3 value stands alone; front vowels are perceived heavily on the value of F1 and the average of F2 and F3 (Ferrand, 2014).

Although specific values can be assigned to each vowel sound at the level of F1, F2 and F3, vowel formant frequencies are not concrete; they vary. Even when spoken in isolation, formant frequencies can differ within a single speaker and they most certainly fluctuate from speaker to speaker. Despite variation, vowels can cause a considerable overlap in formant frequencies when produced by the same speaker or by different speakers. As a result, a one-to-one correspondence between vowels and their formant frequencies is impossible. However, for the average listener, these blurry boundaries between vowels do not typically serve as a complication. Individuals are somehow able to compensate for the differences or similarities in formant frequencies within and between vowels (Ferrand, 2014).

It is evident that a listener with normal hearing does not face extreme challenges when it comes to the perception of vowel sounds based on his or her ability to assign meaning to spoken language from different speakers. Age and gender are two characteristics that cause significant variation in speech and yet for these listeners, such factors have limited effect. A number of theories have been proposed to provide an explanation. Ferrand (2014) suggests that it is not the absolute frequency value that is important, but rather the relationship between and among the formant frequencies. For instances, when the highest vowel sound /i/ is spoken, it is characterized by a very low F1 but a very high F2 and F3, which are close together, following the rule for a front vowel. When an adult male speaker produces this vowel sound, it is very different to the same speech sound produced by a child because of the difference in length of the vocal tract between the two speakers. These two productions, however, maintain a similar relationship between formant frequencies, that is a very low F1 and a very high F2 and F3, closely situated to one another (Ferrand, 2014).

Another theory related to that of formant frequency average stresses the importance of the F2 to F1 ratio. Using the same front vowel, /i/, as an example, Ferrand (2014) explains that with its very low F1 and very high F2, the ratio between the two would be very large. Compare this ratio to that of a back vowel such as /u/, in which F1 and F2 are much closer together. Following these calculations, it was determined that these ratios are very similar despite differences in age and gender. When the ratios of a single vowel sound, /i/, were calculated and compared among men, women, and children, the ratios were very similar; however, when these ratios were examined next to the ratios of a back vowel, /u/, the ratios are very different. This ratio system functions as another way to compare vowels, to notice their distinctions from one another as well as their overlaps with one another (Ferrand, 2014).

While these theories are valuable in understanding how humans perceive vowel sounds, it is not a system that can be applied to every listener in every situation. These techniques assume all listeners have all and equal access to the three formants of each vowel. However, that is not the case for all individuals. People with hearing loss or hearing impairment may or may not have equal access to the three formant frequencies that the average listener is assumed to obtain. Without complete access to all three frequencies, vowel perception becomes unbalanced, meaning some vowels may be easy to hear and understand while others are impossible. It is dependent on what frequencies the listener has access to and this varies significantly from one listener with hearing impairment to the next.

VI. Vowel Perception in Children with Hearing Impairment

Speech perception is critical in the clinical examination of those with hearing impairment because it will provide information regarding what assistive technology is required by the individual. Understanding type, degree, and configuration of the individual's hearing loss will directly translate to the listener's auditory perception as well as phoneme and word recognition (Ferrand, 2014). It is this information that will serve as a reference when determining how to move forward in the clinical sense to meet the needs of the individual with hearing impairment.

The two factors that influence speech perception in individuals with hearing loss are audibility and suprathreshold recognition ability. Audibility addresses the level in which the speech cue is presented in order for the individual with hearing impairment to hear and this level of speech that can be heard by the listener is denoted the suprathreshold level (Ferrand, 2014). In order to accurately perceive speech, individuals with hearing loss must have access to the sound and once that is achieved, their ability to derive meaning from spoken language can be examined.

Deficits are common for any individual with hearing impairment, regardless of age, in discriminating between similar acoustic elements. Therefore, difficulty in identifying the acoustic-phonetic features of speech can be a challenge. For children, as they are in the early stages of language development, hearing loss reduces the amount of acoustic information available to them. They then struggle to piece together these acoustic portions into a single, meaningful message (Ferrand, 2014; Ross, Brackett, & Maxon, 1991). Unlike their normally hearing peers who can associate meaning based exclusively upon audition, children with hearing impairment rely on acoustic-phonetic information to extract the same meaning from the utterance (Ferrand, 2014).

The acoustic structures of a phoneme provide many different cues to perception of speech. As hearing impairments are different to each individual, the cues essential for accurate perception of the utterance depends on the individual and his or her hearing entirely. For some people, the discrimination of vowel can be done with limited issues, while others struggle discerning the subtle differences in formant frequencies that are the entire makeup of a vowel sound. This is a common occurrence for children with high-frequency hearing losses. They often confuse vowels due to lack of frequency resolution and poor audibility. The pure vowels, /i/ and /u/, used in previous examples are two that are difficult to perceive. Since the F1 of each of the vowels are relatively close in frequency (270 Hz for men and 310 Hz for women with /i/; 300 Hz for men and 370 Hz for women with /u/), it is difficult to distinguish which vowel is heard based on that information alone. The F2 for /i/ is relatively high in frequency (between 2300 and 3200 Hz), so children without access to the high-frequencies may not be able to identify the vowel /i/ simply because they cannot hear the acoustic information that would allow for discrimination of this vowel (Ferrand, 2014).

VII. NU-CHIPS

Among the widely accepted speech discrimination tests is the *Northwestern University-Children's Perception of Speech* (NU-CHIPS) test, which is an evaluation developed for use by clinical audiologists in practice (Elliott & Katz, 1980). The test was designed with the pediatric or difficult-to-test population in mind (Dengerink & Bean, 1988). Utilizing monosyllabic nouns that are within the receptive vocabularies of three-year-old children, test takers participate in a closed-response set test that requires a picture-pointing task upon recognition of a spoken word (Dengerink & Bean, 1988).

NU-CHIPS was developed by Elliott and Katz (1980) to study a pediatric listener's perception of 50 monosyllabic words that fall within the receptive vocabulary of young children, determined by the Peabody Picture Vocabulary Test (Dunn and Dunn, 1997; Dobie and Hemel). With a total of 65 words depicted with black-and-white line images, the 15 words not included in the target word population are foils and are said to be phonemically balanced with the target words (Alpiner & McCarthy, 2000; Elliott & Katz, 1980). In a forced-choice identification task, the examinee must point to the correct image within the randomization on a single page, based on what he or she hears from the spoken word of the 50 stimulus pages. A recorded version of the evaluation is available with each test book, providing the option of a male or female speaker (Madell, 1998). The evaluation can be given by the clinician present at time of evaluation as an alternative to the prerecorded tapes provided by the test developers (Elliott & Katz, 1980).

Although each book contains 50 stimulus pages, the evaluation can be administered in whole lists or half-lists, where pages 1 through 25 will provide similar information to pages 26 through 50. The original NU-CHIPS provided two test books, Book A and Book B, utilizing the same monosyllabic words for test words and the same monosyllabic words for foils. Each book was paired with two test forms; Test Form 1 and Test Form 2 correspond with test Book A, while Test Form 3 and Test Form 4 are paired with test Book B (Elliott & Katz, 1980). Since its publication in 1980, test Book B is no longer available in print, although it exists in circulation.

VIII. Purpose of This Study

The purpose of this study is to determine whether the test pages provided by test Book A of the NU-CHIPS diagnostic evaluation were sufficiently similar in the vowel sounds that appeared on each page. According to the manual written by Elliot and Katz (1980), the evaluation consists of 65 phonemically balanced words, all within the receptive vocabulary of three-year-olds. For children with hearing impairment, speech perception is a challenging avenue to explore as not all hearing abilities are the same. In some cases, children do not have access to the first formant, relying only on the second formant to identify the meaning of the spoken word, while other children have access to the first formant and lack that of the second formant, making the first formant the only means for recognition of the vowel and thus, identification of the word. With specific cues required by each child with hearing impairment, the NU-CHIPS evaluation tool was closely analyzed to explore the phonemic balance in terms of vowel sounds used on each test page to therefore determine if it was an appropriate measure of speech perception ability for children with hearing impairment.

Methods

I. Analysis 1: Vowel Difference Analysis

According to test creators Elliott and Katz (1980), the 65 monosyllabic nouns are both within the receptive vocabularies of three-year-old children and phonemically balanced in comparison to one another. To determine the validity of this statement, two separate analyses were conducted in which all 50 test pages of Book A were evaluated. Since Book B is no longer available in print, it could not be analyzed for the study. The four pictorial representations on each page were translated into their corresponding monosyllabic noun, which were then recorded in two separate Excel spreadsheet, one that represented Test Form 1, Test Form 1 – Words, and one that represented Test Form 2, Test Form 2 - Words. When importing the information into the spreadsheet, each page was represented across a single row, where the picture on the top left was placed in column 1, the picture on the top right in column 2, the picture on the bottom right in column 3, and the picture on the bottom right in column 4. This method was maintained across all spreadsheets. For future reference, the target words from Test Form 1 were notated in bold text on Test Form 1 – Words, while the same was done on the second spreadsheet, Test Form 2 – Words.

According to Standard American English, each of the monosyllabic nouns were transcribed into the International Phonetic Alphabet (IPA). These transcriptions were then imported into neighboring Excel spreadsheets named Test Form 1 – Transcriptions and Test Form 2 – Transcription, while still maintaining the bold print target word in IPA for each test form. Furthermore, each word transcription was condensed to the phoneme level; these spreadsheets were duplicated into new spreadsheets where the consonants were eliminated, leaving only the transcribed vowel in Test Form 1 – Vowels and Test Form 2 – Vowels, with the target vowel appearing in bold font.

Once the vowel was identified in its corresponding IPA transcription, the vowels that appeared on each page were compared to one another. For organizational purposes, the 15 transcribed vowel sounds that appeared throughout Book A were color coded for easy differentiation. Beginning with Test Form 1 – Vowels, each page, or row, was independently examined, and the vowels that appeared on the page were categorized as being the same as the target vowel or different from the target vowel. The page was then given a number, either 0, 1, 2, or 3, based on the number of foil vowels that differed from the target vowel, and the number was presented in column 5 of the spreadsheet. These values were then tallied to obtain a total value for how many pages contained foil words with 0 vowels that differed from the target vowel, how many pages contained foil words with 1 vowel that differed from the target vowel, and so on. Once tallied, these numerical values represented the number of foil words with vowels that differed from the target word on each page as well as how many time these differences occurred throughout all 50 pages of Book A, according to the target words presented on Test Form 1. The same procedure was followed for the spreadsheet labeled Test Form 2 – Vowels, where the vowel transcriptions were compared on each page, noting the number that differed from the target vowel, which was then tallied over all 50 pages. These observations seen in Test Form 1 and Test Form 2 are displayed in bar graphs below, Table 1 and Table 2.

II. Analysis 2: Formant Difference Analysis

As a continuation of the first analysis, the second investigation was designed to evaluate the amount of variation in formant frequencies noticed across each individual test page. Recall that different vowels may contain formants with similar frequencies. For examinees with hearing loss at specific frequencies, such vowels that share similar formants may serve as effective foils if the examinees hearing is limited. However, significantly different formant frequencies, or those considered different based upon a percent difference that exceeds 10 percent, determined by researchers prior to the analysis, may offer cues to examinees that could impact their accuracy during the evaluation.

Researchers decided to examine the formant frequencies for two populations of speakers, the adult male and the adult female because the test creators paired Book A of the NU-CHIPS evaluation with two recorded version of test words, one spoken by an adult male and one spoken by an adult female. The formant frequencies obtained for all three formants of each vowel were determined by the extensive vowel study conducted by Peterson and Barney (1952) and these values appear in Appendix I. Continuing the spreadsheet format from analysis one, two additional spreadsheets were created for each test form, one for the adult male speaker and one for the adult female speaker. Test Form 1 – Vowels was copied into the spreadsheet titled Test Form 1 – Formant Frequencies (Men), three separate times to represent F1, F2, and F3. The vowels were then replaced by their numerical formant value for F1, F2, and F3. The same procedure was then followed to create the spreadsheet Test Form 1 – Formant Frequencies (Women) but this time, the values that replaced the vowels corresponded to the adult female speaker determined by Peterson and Barney (1952). For each of the three charts, the target vowel, in terms of formant frequency, appeared in bold print, similar to those in all earlier spreadsheets. Researchers then created spreadsheets, Test Form 2 – Formant Frequencies (Men) and Test Form 2 – Formant Frequencies (Women) by duplicating those of the first test form, with the only difference being the bolded target vowel, in formant frequency, that matched the list provided by Test Form 2.

With the spreadsheet format in place, the formants at each frequency level could be easily compared among the four pictures on a page as they are represented across a single row. At this point, the notation of the target vowel in bold is critical in determining the percent difference of each foil vowel in relation to the target vowel. Each of the 50 pages contains three foil words and one target word; therefore, each page yields three percent differences. The percent difference is calculated by subtracting the foil formant value from the target formant value. This difference is then divided by the target formant value and multiplied by 100 to achieve a percentage. These percentages must be stored in the excel spreadsheet, beside the corresponding row for future reference.

While it is important to stress that the focus of the study is on the word's vowel, a single sound, two exceptions exist. In several cases, the words that appeared on the test page contained either a diphthong or a rhotic vowel. Rather than being categorized as a single sound, these vowels are influenced by more than one sound, where a diphthong is the combination of two vowels, in which the sound begins as one vowel and moves toward another. A rhotic vowel is a vowel followed by the consonant, "r;" although represented by two separate phonetic characters, the sounds are not thought of as being produced individually, but rather blend together to form a single sound, typically resulting in the lowering of the third formant. Since diphthongs and rhotic vowels span a frequency range rather than pinpointing an exact numerical value for each formant, the test pages that utilized either one as the target vowel was eliminated from the study. In the case of a diphthong or rhotic vowel appearing as a foil rather than the target word, researchers automatically considered the vowel to exceed the 10 percent rule that was established at the start of the project, because of the variation that would be seen across the frequency range.

After the consideration of diphthongs and rhotic vowels, nine pages (11, 12, 14, 16, 21, 36, 42, 44, and 47) were eliminated from the study, leaving 41 examinable pages. Based on the mathematical process above, three percent differences were calculated for each of the 41 pages for all three formant frequencies, F1, F2, and F3, for the average adult male speaker as well as the average adult female speaker for both Test Form 1 as well as Test Form 2 and recorded next to their respective tables. Once all percentages were computed, those that consistently showed a zero percent difference, meaning that all pages that shared the same vowel across all four image labels were eliminated. Since researchers were only concerned with the difference among foil vowels and target vowels, these were no longer necessary during the evaluation. For the remaining 20 pages, three different formats existed; one foil word could vary from the target while the other two maintained the target vowel; two vowels could vary from the target, with the third foil word including the same vowel as the target word, or all three foil words could have vowels that differed from the vowel in the target word. All three test page scenarios were considered independently and the number of percent differences that exceeded the rule established by researchers, 10 percent, were tallied for both male and female speakers in Test Form 1. The results were converted into a visual graph for each speaker. This same procedure was followed for Test Form 2, producing two more bar graphs displayed below. Any vowel differences between target and foil words that led to a percent difference that failed to exceed the 10 percent rule was considered sufficiently similar and no further evaluation was necessary.

Results

I. Analysis 1: Vowel Difference Analysis

Of the 50 test pages of Book A of the NU-CHIPS evaluation tool that correlate to the test words provided by Test Form 1, 21 were found to contain zero foil words with vowels that differed from the target word. In other words, all foil words were comprised of the same vowel that was observed in the target word. Test pages that differed from the target vowel in 1, 2, or 3 of the foil words occurred in 17, 11, and 1, cases respectively. These results are shown in Table 1.

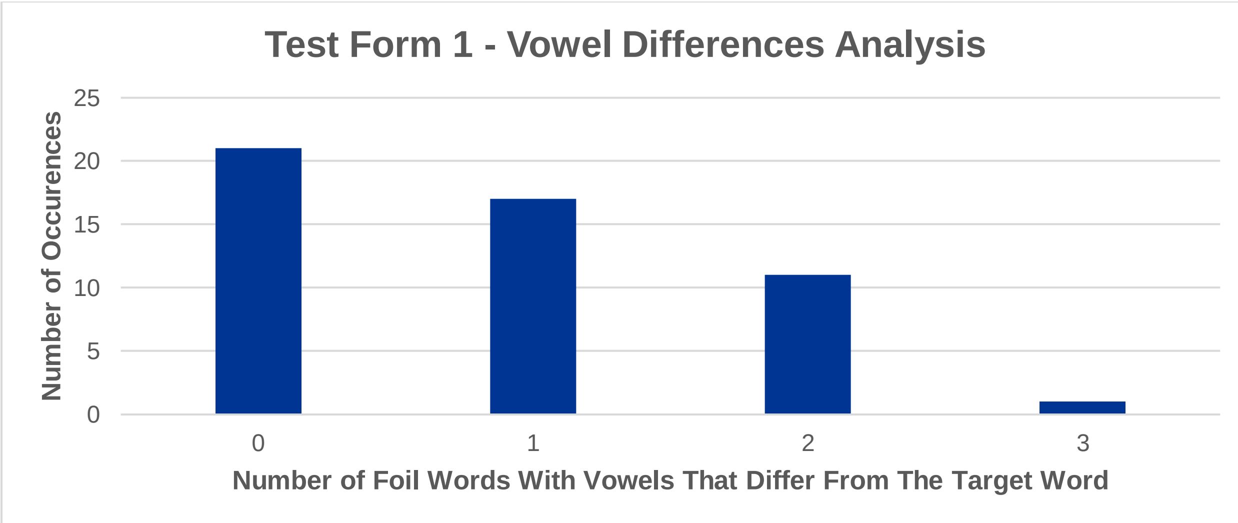

Table 1: Test Form 1 – Vowel Difference Analysis

Of the 50 test pages of Book A of the NU-CHIPS evaluation tool that correspond to the test words provided by Test Form 2, 21 were found to contain zero foil words with vowels that differed from the target word. In other words, all foil words were comprised of the same vowel that was observed in the target word. Test pages that differed from the target vowel in 1, 2, or 3 of the foil words occurred in 17, 9, and 3, cases respectively. These results are shown in Table 2.

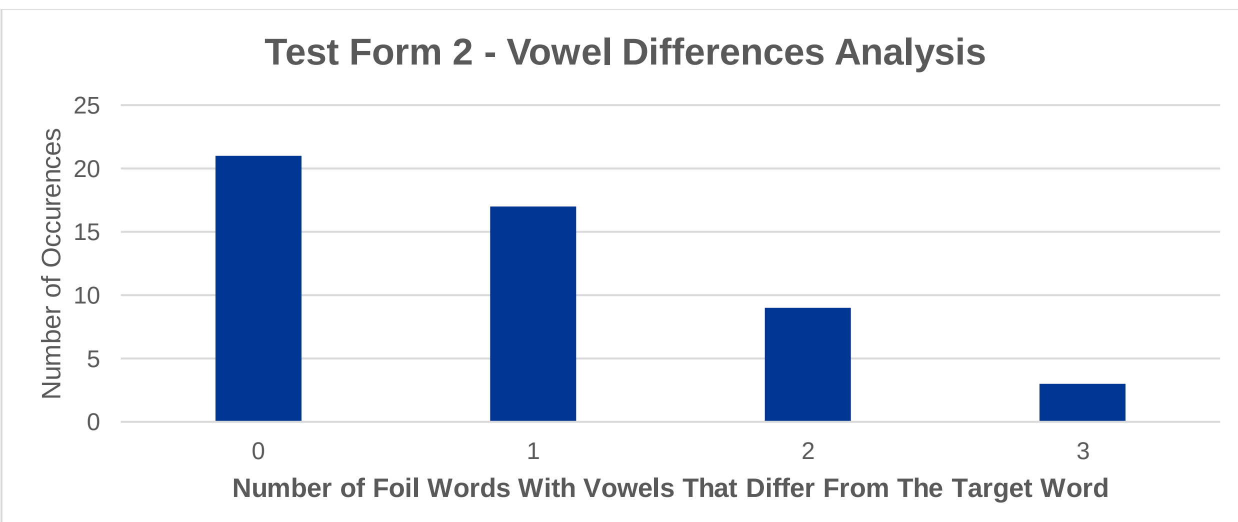

Table 2: Test Form 2 – Vowel Difference Analysis

II. Analysis 2: Formant Difference Analysis

In Book A of the NU-CHIPS evaluation tool, 41 of the 50 pages fit the criteria established by researchers and were studied during analysis one. Of these 41 pages, 20 contained noticeable percent differences in vowel variation; the other 21 showed a zero percent difference when comparing the vowel of the target word to the vowel of the foil words. These pages with similar vowels were eliminated from further investigation.

According to researchers, a foil vowel was considered *different* from the target vowel when the percent difference between the two exceeded the predetermined 10 percent rule. When analyzing the formant frequencies, F1, F2, and F3, of the average adult male speaker, a test page that varied by only one foil vowel saw 10 foil vowels that were considered *different* for F1, 13 for F2, and 11 for F3. For test pages that contained two foil vowels that varied from the target vowel, 9 were marked as *different* from the target vowel for F1, 7 for F2, and 2 for F3. When a page saw a variation across all three foil vowels, 1 foil vowel exceeded the 10 percent rule for F1, 0 for F2, and 0 for F3. These results are shown in Table 3.

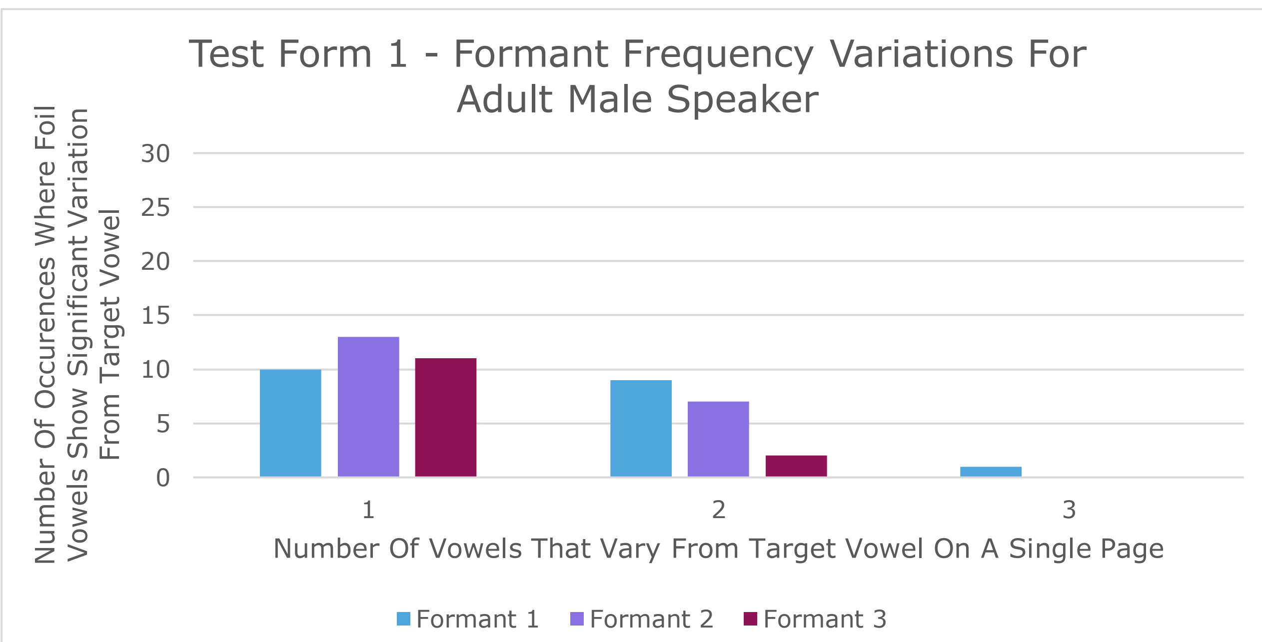

Table 3: Test Form 1 – Formant Frequency Variations for Adult Male Speaker

Following the same procedure detailed above, the formant frequencies of the average adult female speaker were condensed into a graphical display that compared the percent differences between the foil vowels and the target vowel on the 20 pages that saw variation. When the test page encompassed one vowel difference, 10 foil vowels were labeled different for F1, 10 for F2, and 11 for F3. Of the pages that saw two separate foil vowels, 9 different from the target for F1, 9 for F2, and 1 for F3. For the pages in which all vowels were unique, or three vowels differed from the target vowel, 1 foil vowel was determined different for F1, 1 for F2, and 0 for F3. Results can be observed in the bar graph presented in Table 4.

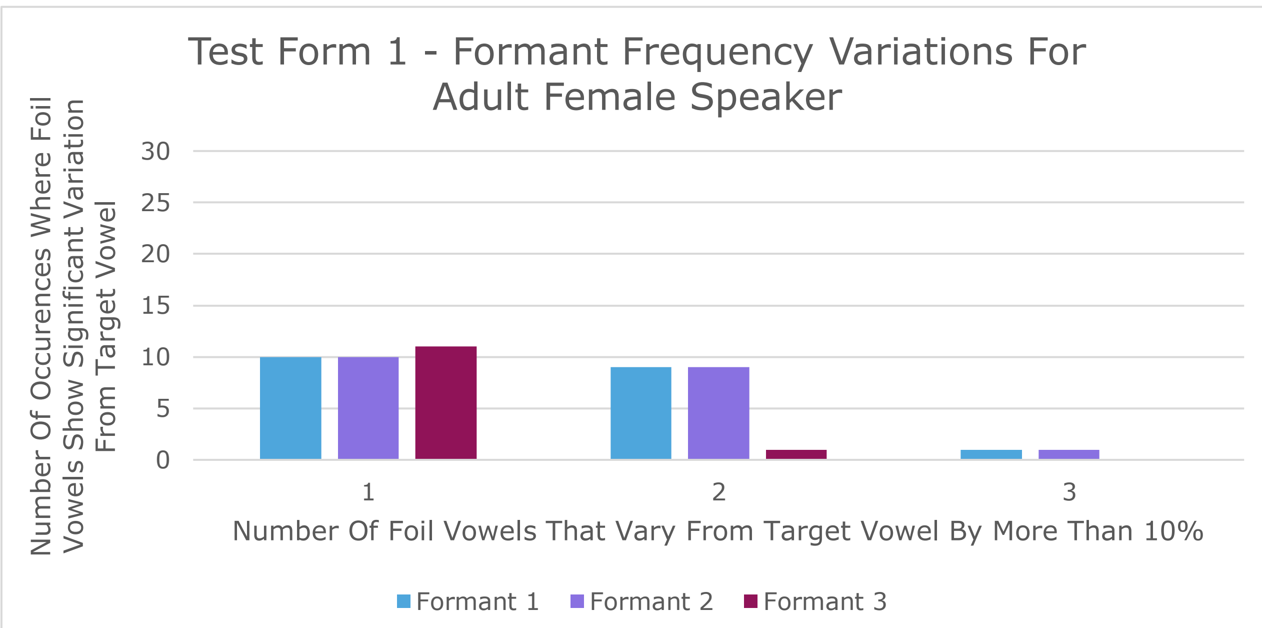

Table 4: Test Form 1 – Formant Frequency Variations for Adult Female Speaker

Test Form 2 experienced an identical evaluation to Test Form 1. For the average adult male speaker, those pages that varied by one foil vowel, while the other two foil vowels matched that of the target word, 10 exceeded the 10 percent rule for F1, 11 for F2, and 12 for F3. When the test page revealed two vowels that varied from the target, 7 were deemed *different* for F1, 7 for F2, and 2 for F3. When all, three, foil vowels were dissimilar to the target, 3 were considered *different* in terms of F1, 2 for F2, and zero for F3. All of these results presented in Table 5 were based off of the 20 pages that fell within the examination capacity of researchers.

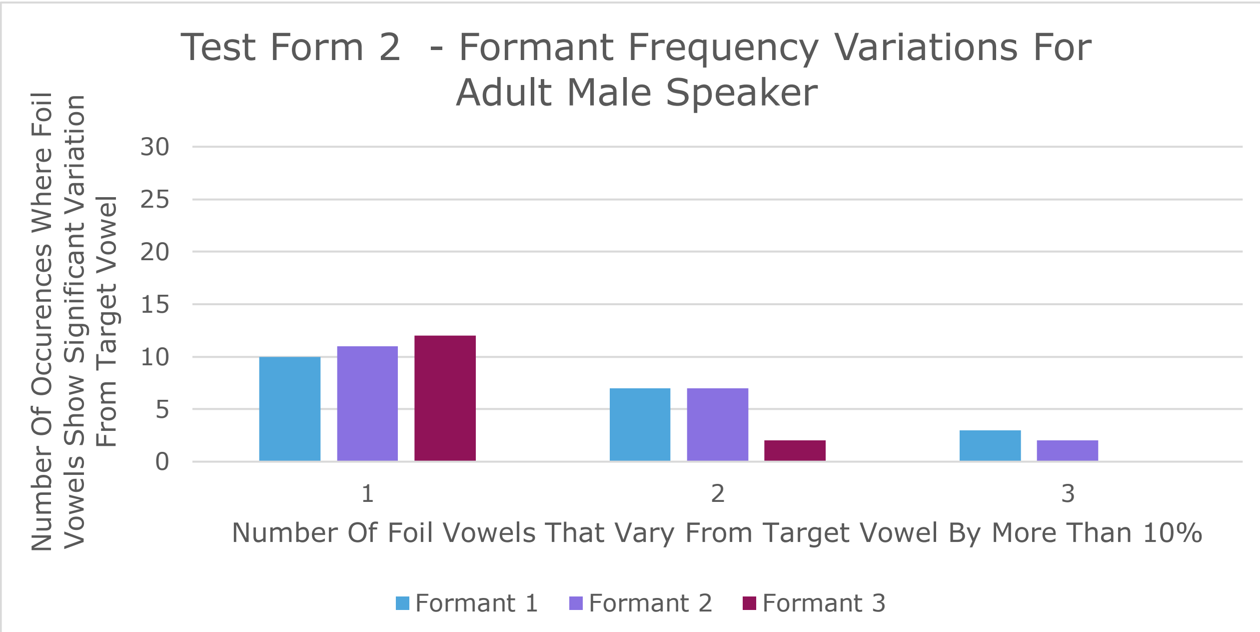

Table 5: Test Form 2 – Formant Frequency Variations for Adult Male Speaker

Lastly, Test Form 2 was analyzed further to account for the average female speaker's formant frequencies. In scenario one, where only one vowel stood apart from the others, 10 vowels were termed *different* for F1, 10 for F2, and 13 for F3. Those pages that saw two separate vowels, one target vowel that appeared in two pictures, and one foil vowel shown in the remaining two images, 7 vowels varied by 10 percent or more in F1, 7 for F2, and 1 for F3. When all vowels varied from the target, 3 were categorized as *different* compared to the target in terms of F1, 3 for F2, and 0 for F3. Evidence for these calculations are outlined in the bar graph below, Table 6.

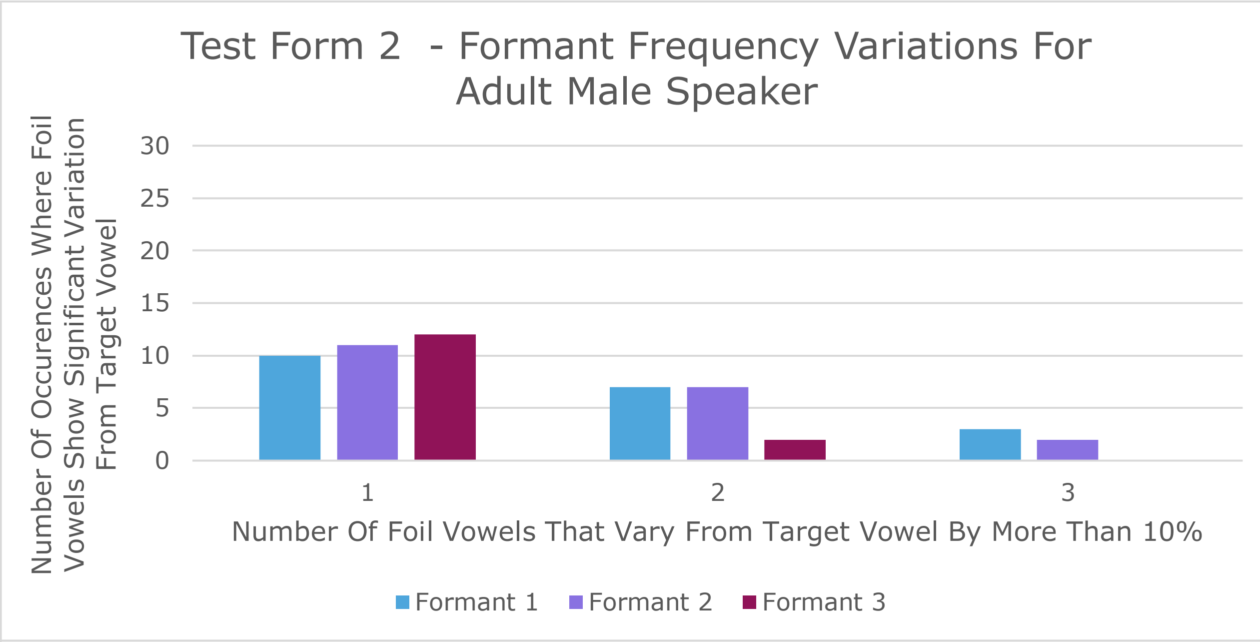

Table 6: Test Form 2 – Formant Frequency Variations for Adult Female Speaker

Discussion

I. Analysis 1: Vowel Difference Analysis

The results displayed in the bar graphs in Analysis 1, Tables 1 and 2, show the number of foil words that contain a vowel that differ from the vowel within the monosyllabic target word provided on each test page. On 21 of the 50 pages, none of the vowels varied; all four vowels were identical, which led researchers to believe that the examinee's consonant perception could be determined from these pages in both Test Form 1 and Test Form 2.

For the remaining 29 test pages, however, vowel consistency was not present. Of these pages that accompanied the test word list of Test Form 1, 17 pages varied by one vowel, 11 pages varied by two vowels, and 1 page varied by three vowels. Since a unique target word list was generated for Test Form 2, despite still utilizing test Book A, slightly different values were noted where 17 pages varied by one vowel, 9 pages varied by two vowels, and on 3 of the pages, three vowels varied.

These values call attention to the vowel variation in the NU-CHIPS evaluation. As previously stated, when the vowels were consistent across a page, the examinee's consonant perception could be observed, but when the vowels on a page vary by one, two, or three vowels, the same assumption of consonant perception cannot be applied. While consonants were not at the focus of this study, researchers noticed that in the cases of vowel variation, the consonants presented across the page were inconsistent, meaning vowel perception could not be determined either. When these differences exist, the purpose intended by test developers becomes unclear because the independent variable has now been eliminated.

Without any recognition of these differences in the manual provided by test developers, researchers were concerned with what this could mean for children with hearing loss who make up the population of examinees. If significant enough, the vowel variation could cause examinees with hearing loss to automatically eliminate a foil word based on the vowel alone, automatically increasing their chance of accuracy in selecting the target word from the group of four pictures. When one vowel varied from the target, if significant enough, the child may have a one in three chance of choosing the correct answer rather than a one in four chance and thus, the probability increases as the number of vowels that vary rise across a page so that on the pages where four different vowels exist, the vowel's cues alone may give more information based on the child's hearing ability that was unaccounted for by NU-CHIPS developers. While this is the most extreme case and happens very few times throughout the evaluation, this inconsistency is present and will have an influence on the examinee's results.

II. Analysis 2: Formant Difference Analysis

With knowledge of how the vowel in each of the foil words vary in comparison to the vowel in the target word, researchers wanted extend this study to locate the exact difference between varying vowels on a single page. To do this, percent differences were calculated between the formant values of all foil words and their corresponding target word, at the level of F1, F2, and F3, for both test forms and for the two different speakers. As determined by anaylsis one, 21 pages utilized the same vowel in all four pages and thus made up the ideal pages of the NU-CHIPS evaluation.

In conjunction with analysis one, analysis two can provide a significant amount of information for researchers. The vowel difference graph in Table 1 parallels the formant frequency variation in Table 3 and Table 4 for both the average male and the average female speakers; the vowel differences graph in Table 2 relate to the formant frequency variations in Table 5 and Table 6 for both speakers. The first set of data is reflective of the word list in Test Form 1 and the second group of corresponding charts coincide with the word list described in Test Form 2.

After close examination of the information presented in both analyses, researchers determined that for both test forms, when one vowel varied on a page, that vowel varied by more than 10 percent for F1 on the majority of test pages across both test forms and both speakers. For children with high-frequency hearing loss, who may only have access to the lower frequencies, this significantly *different* vowel increases his or her chances of selecting the correct image. Instead of the child having a 25 percent chance of accuracy, he or she may raise his or her chances to 33.34 percent based on the vowel alone, with no indication of the influence of the consonants.

When two of the vowels on the test page are distinguishable from the target vowel, the majority of pages affected saw F1 and F2 frequency formant percent difference that exceeded 10 percent. However, the higher frequencies, F3, saw significantly less variation between the foils and the target, meaning they were much more closely related and therefore, could not be discerned on the basis of vowel frequency alone. This was the case in both Test Form 1 and Test Form 2 for both the adult male speaker as well as the adult female speaker. Similar to the observation made when only one foil word appeared to have a varying vowel, a child with access to the lower frequencies, due to a high-frequency hearing loss, may be able to automatically eliminate those two foil words because of their significantly different vowel sound. Now, the examinee's chances are even higher because they can disregard half of their choices, providing them with a 50 percent chance of selecting the desired target word determined by test developers.

Although very few pages displayed four different vowels, or three that differed from the target vowel, this scenario existed on 1 page in Test Form 1 and 3 pages in Test Form 2. For Test Form 1, the affected page contained all three foil vowels that varied by more than 10 percent at the F1 level, leading researchers to regard this as a significant difference. Based on this evidence, the examinee with a high-frequency hearing loss may be able to select the correct answer based on the information provided by F1 and therefore, it does not measure the child's speech perception ability, but rather what the child can hear. This was true of the first formant of the male speaker of Test Form 1, but with the female speaker, this same idea applied to F2 as well, where all three foil vowels varied by more than 10 percent of the target vowel.

In terms of Test Form 2, more pages existed that were made up of three different vowels in all of the foil words. At the first formant frequency level for the average adult male speaker, all 3 pages exceeded the 10 percent rule, and 2 of the 3 exceeded the rule at the second formant frequency level. Instead of matching the picture based on the entirety of the spoken word, the examinee may be able to identify the target word exclusively because of the vowel variation on 2 or 3 of the pages. With the female speaker, all 3 vowels differed in a significant manner, exceeding the 10 percent mark, at both the F1 and F2 level. For both formant frequency levels, 6 percent of the test is now affected and these test pages no longer match the intentions of test developers, which can have a large influence when the evaluation is scored.

III. Limitations

This study only explored the NU-CHIPS evaluation in terms of the vowel sounds presented in all 65 of the monosyllabic words presented by test developers. The area of consonants was not within the scope of research. Since consonants provide their own cues, independent and unique to those of vowels, they are likely to impact how a child perceives a monosyllabic word. Not only do they have their own frequencies and indicative features, but consonants also affect the vowels they precede and follow due to coarticulation, which is an area that could expanded upon with further research.

The majority of pages in test Book A fit the criteria for the study, but those that did not could not be investigated completely. These pages were the ones that contained a diphthong or a rhotic vowel in the target word; they failed to be included because their exact formant frequency are not accurately displayed with a numerical value like the monophthongs. Additionally, these vowels were considered *different* when they appeared as a foil vowel because of their formant transitions, which tend to show a wide range of variation in frequencies throughout the vowel sound. Without the exact values, the difference percentages could not be properly calculated, but with continued investigation, these values may provide further insight into the evaluation.

Additionally, this study reflects the speech perception of monosyllabic words and is not representative of speech perception in other contexts, such as connected speech. This test serves as a measure of auditory perceptual ability but cannot be directly extended into multisyllabic words, multiword phrases, or complete sentences. Rather than being representative of all speech encounters, the examination provides information on what an examinee can hear based on how he or she derives meaning from the single word presented during the evaluation.

Researchers focused on what was provided by the test developers to evaluate the speech perception test, which resulted in the examination of the average formant frequencies of an adult male speaker as well as the average formant frequencies of an adult female speaker. This provided a baseline for the study, however, it is unclear how many audiologists use the recording when giving this test. When recordings are not used, the information presented in this study may not apply due to the dialectical variation that exists throughout the country. With this in mind, the study could be extended to determine how this test is used with practicing audiologists, which would provide even more information about the application of the test tool.

V. Summary

Based on the findings of this study, the majority of pages presented in Book A of the NU-CHIPS evaluation contained vowels that varied in comparison with the target vowel; when observed in relation to the formant frequencies, several pages were deemed insufficiently similar because the variation between the foil vowels and the target vowel exceeded the criteria determined by the researchers. For children with hearing impairment, especially those without access to the higher frequencies, Test Form 1 and Test Form 2 did not appear to be phonemically balance between foil words and target words, based exclusively on vowel sounds. This may influence the examinee’s scores because the variation is significant enough for foils to be eliminated in the closed-set examination, increasing the chance of accuracy in selecting the target word from the foil words on a number of pages. The test population must be considered prior to using the NU-CHIPS evaluation as it may not be the most appropriate measure of speech perception abilities for children with hearing impairment due to these vowel variations.

Appendix I

Men

Front Vowels	Formant 1 (F1)	Formant 2 (F2)	Formant 3 (F3)
/i/	270	2290	3010
/ɪ/	390	1990	2550
/ɛ/	530	1840	2480
/æ/	660	1720	2410

Back Vowels	Formant 1 (F1)	Formant 2 (F2)	Formant 3 (F3)
/u/	300	870	2240
/ʊ/	440	1020	2240
/o/	570	840	2410
/ɑ/	730	2290	2440

Central Vowels	Formant 1 (F1)	Formant 2 (F2)	Formant 3 (F3)
/ʌ/	640	1190	2390
/ɝ/	490	1350	1690

Women

Front Vowels	Formant 1 (F1)	Formant 2 (F2)	Formant 3 (F3)
/i/	310	2790	3310
/ɪ/	430	2480	3070
/ɛ/	610	2330	2990
/æ/	860	2050	2850

Back Vowels	Formant 1 (F1)	Formant 2 (F2)	Formant 3 (F3)
/u/	370	950	2670
/ʊ/	470	1160	2680
/o/	590	920	2710
/ɑ/	850	1120	2810

Central Vowels	Formant 1 (F1)	Formant 2 (F2)	Formant 3 (F3)
/ʌ/	760	1400	2780
/ɝ/	500	1640	1960

Derived from “Control Methods Used in a Study of the Vowels” (Peterson & Barney, 1952).

References

Alpiner, J. G., & McCarthy, P. A. (2000). Rehabilitative Audiology Children and Adults (3rd ed.). Baltimore, MD: Lippincott Williams & Wilkins.

Cienkowski, K. M., Ross, M., & Lerman, J. (2009). The Word Intelligibility by Picture Identification (WIPI) Test Revisited. *Journal of Educational Audiology*, 39-43. Retrieved February 02, 2017, from http://www.edaud.org/journal/2009/4-article-09.pdf

Dengerink, J.E., & Bean, R. E. (1988). Spontaneous Labeling of Pictures on the WIPI and NU-CHIPS by 5-Year-Olds. *Language Speech and Hearing Services in School, 19*(2), 144. doi:10.1044/0161-1461.1902.144

Dobie, R.A., & Hemel, S.V. (Eds.). (2004). Hearing Loss: Determining Eligibility for Social Security Benefits. Washington, DC: National Academies Press.

Dunn, L.M., & Dunn, L.M., (1997) Peabody Picture Vocabulary Test, Third Edition. Circle Pines, MN: American Guidance Service.

Eisenberg, L. S., Johnson, K. C., & Martinez, A. S. (2005, September 12). Clinical Assessment of Speech Perception for Infants and Toddlers. Retrieved Februrary 26, 2017, from http://www.audiologyonline.com/articles/clinical-assessment-speech-perception-for-1016

Elliott, L. L., Katz, D. R., (1980). Northwestern University-Children's Perception of Speech: Technical Manual. St. Louis, MO: AUDITEC of St. Louis.

Ferrand, C. T. (2014). Speech Science an Integrated Approach to Theory and Clinical Practice (3rd ed.). Upper Saddle River, NJ: Pearson Education, Inc.

Justice, L. M., & Redle, E. E. (2014). Communication Sciences and Disorders a Clinical Evidence-Based Approach (3rd ed.). Upper Saddle River, NJ: Pearson Education, Inc.

Madell, J. R. (1998). Behavioral Evaluation of Hearing in Infants and Young Children. New York: Thieme.

Martin, F. N., & Clark, J. G. (2009). Introduction to Audiology (10th ed.). Boston, MA: Pearson Education, Inc.

National Institute on Deafness and Other Communication Disorders (NIDCD). (2017, February 10). Quick Statistics About Hearing. Retrieved March 27, 2017, from https://www.nidcd.nih.gov/health/statistics/quick-statistics-hearing

Northern, J.L., & Downs, M.P. (2002). Hearing in Children (5th ed.). Baltimore, MD: Lippincott Williams & Wilkins.

Peterson, G.E., & Barney, H.L. (1952). Control Methods Used in a Study of the Vowels. *The Journal of the Acoustical Society of American, 24*(2), 175-184. doi:10.1121/1.1906875

Raphael, L.J., Borden, G.J., & Harris, K.S. (2007). Speech Science Primer: Physiology, Acoustics, and Perception of Speech (5th ed.). Baltimore, MD: Lippincott Williams & Wilkins.

Ross, M., Brackett, D., & Maxon, A.B. (1991). Assessment and management of mainstreamed hearing-impaired children. Austin, TX: Pro-Ed.

Shargorodsky, J., Curhan, S., Curhan, G., & Eavey, R. (2010) Change in Prevalence of Hearing Loss in US Adolescents. *Journal of the American Medical Association, 304*, 772-778.

Small, L. H. (2012). Fundamentals of Phonetics A Practical Guide for Students (3rd ed.). Upper Saddle River, NJ: Pearson Education, Inc.

Strange, W. (1999). Perception of Vowels: Dynamic Consistency. In J.M. Pickett (Ed.), *The acoustics of speech communication: Fundamentals, speech perception in theory, and technology.* (pp. 153-165). Boston: Allyn and Bacon.

Tatham, M., & Morton, K. (2011). A guide to speech production and perception. Edinburgh University Press.