

# **HHS Public Access**

Author manuscript

J Am Acad Audiol. Author manuscript; available in PMC 2018 November 01.

Published in final edited form as:

J Am Acad Audiol. 2017 ; 28(10): 901–912. doi:10.3766/jaaa.16105.

## **Segmental and Suprasegmental Perception in Children Using Hearing Aids**

## **Kaitlyn A. Wenrich**1, **Lisa S. Davidson**1,2,3, and **Rosalie M. Uchanski**1,2

<sup>1</sup>Program in Audiology and Communication Science, Washington University School of Medicine, St. Louis, MO

<sup>2</sup>Department of Otolaryngology, Washington University School of Medicine, St. Louis, MO

<sup>3</sup>Central Institute for the Deaf, St. Louis, MO

## **Abstract**

**Background—**Suprasegmental perception (perception of stress, intonation, "how something is said" and "who says it") and segmental speech perception (perception of individual phonemes or perception of "what is said") are perceptual abilities that provide the foundation for the development of spoken language and effective communication. While there are numerous studies examining segmental perception in children with hearing aids (HAs), there are far fewer studies examining suprasegmental perception especially for children with greater degrees of residual hearing. Examining the relation between acoustic hearing thresholds, and both segmental and suprasegmental perception for children with HAs may ultimately enable better device recommendations (bilateral HAs, bimodal devices [one CI and one HA in opposite ears], bilateral CIs) for a particular degree of residual hearing. Examining both types of speech perception is important because segmental and suprasegmental cues are affected differentially by the type of hearing device(s) used (i.e., cochlear implant [CI] and/or hearing aid [HA]). Additionally, suprathreshold measures, such as frequency resolution ability, may partially predict benefit from amplification and may assist audiologists in making hearing device recommendations.

**Purpose—The purpose of this study is to explore the relationship between audibility (via hearing** thresholds and speech intelligibility indices), and segmental and suprasegmental speech perception for children with hearing aids. A secondary goal is to explore the relationships amongst frequency resolution ability (via spectral modulation detection measures), segmental and suprasegmental speech perception, and receptive language in these same children.

**Research Design—**A prospective cross-sectional design.

Correspondence: Kaitlyn A. Wenrich, Washington University School of Medicine in St. Louis, Program in Audiology and Communication Science, 660 S. Euclid Ave, Campus Box 8115, St. Louis, MO 63110, katoner@wustl.edu, (610)-301-5430. **Conflict of Interest Statement:** We have no conflicts of interest to report.

**Presentations:** Portions of this research were presented at the Association for Clinical and Translational Science 'Translational Science 2015' Meeting on April 16–18, 2015 in Washington, D.C. and the Mid South Conference on Communication Disorders on February 25–26, 2016 in Memphis, TN.

**Study Sample—**Twenty-three children, ages 4 years 11 months to 11 years 11 months, participated in the study. Participants were recruited from pediatric clinic populations, oral schools for the deaf, and mainstream schools.

**Data Collection and Analysis—**Audiological history and hearing device information were collected from participants and their families. Segmental and suprasegmental speech perception, spectral modulation detection, and receptive vocabulary skills were assessed. Correlations were calculated to examine the significance  $(p<0.05)$  of relations between audibility and outcome measures.

**Results—**Measures of audibility and segmental speech perception are not significantly correlated, while low-frequency PTA (unaided) is significantly correlated with suprasegmental speech perception. Spectral modulation detection is significantly correlated with all measures (measures of audibility, segmental and suprasegmental perception and vocabulary). Lastly, although age is not significantly correlated with measures of audibility, it is significantly correlated with all other outcome measures.

**Conclusions—**The absence of a significant correlation between audibility and segmental speech perception might be attributed to overall audibility being maximized through well-fit hearing aids. The significant correlation between low-frequency unaided audibility and suprasegmental measures is likely due to the strong, predominantly low-frequency nature of suprasegmental acoustic properties. Frequency resolution ability, via spectral modulation detection performance, is significantly correlated with all outcomes and requires further investigation; its significant correlation with vocabulary suggests that linguistic ability may be partially related to frequency resolution ability. Last, all of the outcome measures are significantly correlated with age, suggestive of developmental effects.

#### **Keywords**

hearing aids; hearing loss; pediatric; speech perception; suprasegmental; spectral resolution

## **INTRODUCTION**

For non-tonal languages, speech perception skills may be broadly categorized into two main areas; segmental perception and suprasegmental perception. Segmental perception requires encoding of individual phonemes (vowels and consonants) of the speech signal to facilitate understanding words and sentences. Suprasegmental perception requires encoding of intonation, stress, and rhythm in the speech signal to facilitate discerning specific attributes of the talker such as talker identity, age, dialect and emotion. Segmental perception may be viewed as understanding the meaning of speech or "what is said" and suprasegmental perception as "how it is said." Both types of information are simultaneously encoded in the speech signal and are necessary for effective spoken communication (Abercrombie, 1967; Pisoni, 1997).

For typically developing infants with normal hearing sensitivity (NH), both segmental and suprasegmental perception play critical roles in the development of spoken language. For example, infants with NH presumably attend to the acoustic cues of intonation, stress and rhythm (suprasegmental perception; information regarding how something is said or who

said it) before they are able to interpret phonemes unique to their spoken language (Werker and Yeung, 2005). And, for spoken English, infants seemingly attend to the typical intonation patterns of natural speech and to the exaggerated intonation patterns of childdirected speech to identify word boundaries which enable the learning of words (Jusczyk et al, 1999; Seidl and Cristia, 2008; Seidl and Johnson, 2008; Swingley, 2009). These abilities, to hear intonation patterns and then attend to them, are thought to be fundamental to early language processing and vocabulary acquisition (Thiessen et al, 2005; Werker and Yeung, 2005; Massicotte-Laforge and Shi, 2015).

The presence of hearing loss (HL) at birth or during early infancy disrupts these critical perceptual processes (segmental and suprasegmental) resulting in well documented deficits in spoken language skills (Lederberg et al, 2000). Yet current standard clinical protocols focus primarily on assessing segmental properties of speech perception, such as phoneme, word and sentence perception (American Academy of Audiology, 2013). There are many recent reports on segmental speech perception abilities for children with varying degrees of HL and various hearing devices (i.e., HAs and/or CIs) (Dunn et al, 2014; Walker et al, 2015; Wolfe et al, 2015).

Interestingly, several decades ago prior to the wide spread use of CIs, both segmental and suprasegmental perception skills were often examined in children with HL (Hack and Erber, 1982; Boothroyd, 1984). Both types of perception were examined in an effort to understand how amplified speech could combine with visual speech cues to facilitate comprehension of spoken syllables and words. Most of the children in these studies had severe to profound HL and most were tested under earphones, i.e., they were not tested with their personal HAs. Erber and colleagues conducted a number of studies that established a relationship between hearing level (PTA at .5, 1 and 2 kHz) and perception of syllable stress and of speech features in children ages 9–15 years old (Erber 1972a; Erber 1972b; Erber 1979; Hack and Erber 1982). They found, overall, that perception of stress pattern (a suprasegmental aspect of speech) was better than perception of speech segments or speech features. Similarly, Boothroyd conducted experiments that assessed suprasegmental and segmental speech perception in children with HL (11–19 years old). Listeners were grouped by PTA for subsequent analyses. Overall, group performance decreased with increasing HL, and performance was better on most of the suprasegmental tests than on most of the segmental tests (exceptions: perception of vowel height was better than of talker-gender and perception of intonation was poorer than perception of consonant voicing and continuance). Although these decades-old studies with children with HL evaluated both segmental and suprasegmental perception, their results are neither definitive nor comprehensive. For example, some questions left unanswered are: How would these children perform using their HAs, or modern HAs, or currently available CIs? How would children with less-severe HL perform? How would children with HL perform on other tests of suprasegmental perception, such as discrimination amongst female talkers or discrimination of syllable stress?

More recently, a number of studies have reported on segmental and suprasegmental speech perception in children with HL, especially those with severe-to-profound HL who use CIs or HAs (Most and Peled, 2007; Geers et al, 2013). This interest in examining both types of speech perception is motivated by clinicians' desire to recommend the best devices for a

child, including the option of bimodal devices (one CI and one HA in opposite ears), and by the knowledge that segmental and suprasegmental cues may be affected differentially by the type of hearing device(s) (i.e., CI and/or HA) (Kong et al, 2005). While children with CIs typically perform better on segmental perception (e.g., word and sentence recognition) than children with severe-to-profound losses with HAs, the opposite is true for suprasegmental perception (e.g., discrimination of stress, pitch and emotion) (Osberger et al, 1991; Boothroyd and Eran, 1994; Blamey et al, 2001; Most and Peled 2007; Most et al, 2011). HAs can convey suprasegmental information well, but are limited, primarily by the degree of HL of the listener, in the delivery of the spectral content of speech segments. By contrast, CIs do not convey very well the low frequency, voice-pitch-related acoustic information needed for suprasegmental perception (Grant, 1987; Carroll and Zeng, 2007; Most and Peled, 2007; Nittrouer and Chapman, 2009; Most et al, 2011), but do convey fairly well temporal-envelope information across a wide range of frequency regions such that recognition of vowels and consonants (segmental perception) is often very good (Van Tasell et al, 1992; Shannon et al, 1995).

For listeners with HL who use bimodal devices, most research suggests that low frequency acoustic cues transmitted via the HA may be responsible for their good speech perception abilities (Tyler et al, 2002; Ching et al, 2007; Most and Peled, 2007; Dorman et al, 2008; Nittrouer and Chapman 2009; Most et al, 2011; Sheffield and Zeng, 2012). However, results are varied as to the degree of residual hearing necessary for bimodal benefit and most studies are from adult participants. One study with adults (Dorman et al, 2015) suggests that thresholds < 60 dB HL at 125–500 Hz predict maximum benefit for speech recognition in noise while another suggests that thresholds at frequencies < 500 Hz should not exceed 80 dB HL (Illg et al, 2014). In addition to predicting bimodal benefit from acoustic hearing thresholds, other studies with adults examined the utility of frequency resolution ability (spectral modulation detection, SMD, or spectral ripple perception) in predicting bimodal benefit. Zhang and colleagues (2013) found a relationship between bimodal benefit and the low-frequency hearing thresholds of the HA ear, the aided speech perception ability of the HA ear and the SMD threshold at the HA ear. When the adult listeners were divided by HL (mild-moderate low-frequency HL vs. severe-profound low-frequency HL), only the SMD threshold was significantly correlated with bimodal benefit within both groups. In several other studies with adults, frequency resolution ability (SMD or spectral ripple perception) was reported to be significantly correlated with segmental speech perception (Litvak et al, 2007; Won et al, 2007; Saoji et al, 2009; Spahr et al, 2011). Thus, assessing a listener's frequency resolving ability has the potential to assist clinicians in determining a listener's benefit from using a HA, especially when combined with a CI (Bernstein et al, 2013). Additionally, for pediatric listeners, the utility of examining frequency resolution ability can be particularly valuable since frequency-resolution tasks, such as SMD, are thought to be non-linguistic. Eliminating or minimizing confounds attributable to children's language abilities (Rakita, 2012; Bernstein et al, 2013) is always desirable, especially when considering the utility of these measures for making device recommendations for very young children.

For pediatric HA users, several recent studies have examined relations between hearing levels, audibility, speech perception, and language skills (Blamey et al, 2001; Davidson and

Skinner, 2006; Scollie, 2008; Stiles et al, 2012). In these studies, hearing levels were represented by PTAs and aided audibility was represented by the Speech Intelligibility Index (SII). Strong relations were found, between 'hearing' and both speech perception and language skills. Again, however, only segmental speech perception, not suprasegmental perception was examined.

A better understanding of the link between hearing level and perceptual outcomes is critical for pediatric patients who, because of their inchoate cognitive and language levels, are unable to complete extensive speech perception assessments. And, yet, since CI candidacy is frequently determined by 12 months of age or younger, the audiological criteria for these very young children are based primarily on degree of HL (i.e., PTA). And, such audiological criteria for CI candidacy are based on empirical studies of segmental-type speech perception abilities of children using CIs or HAs, but not both types of devices together (Boothroyd and Eran, 1994; Eisenberg et al, 2000; Davidson and Skinner, 2006). Examining both suprasegmental and segmental perception may help researchers develop an understanding of hearing loss' effects across the continuum of severity of HL. An improved understanding would be especially important for hearing device recommendations that are transitions (i.e., 2 HAs to bimodal devices, bimodal devices to 2 CIs). Clinicians should know how much acoustic hearing is enough to provide both good segmental and suprasegmental perception.

The long term objectives of our research program are to understand better how degree of HL relates to segmental and suprasegmental perception, and how acoustic hearing contributes to bimodal benefit for spoken language for children with HL who are developing spoken language skills. This current study examines speech perception, both segmental and suprasegmental, afforded by acoustic hearing alone. The goal of this report is to determine the relationships, for pediatric hearing aid users, amongst audibility (via hearing thresholds and speech intelligibility indices [SII]), segmental speech perception, and suprasegmental speech perception. A secondary goal is to explore the relationships amongst frequency resolution ability, segmental and suprasegmental speech perception, and receptive language in this same population.

#### **METHODS**

#### **Design**

This was a prospective, cross-sectional study. An audiologic evaluation was completed with all participants and their personal hearing aids were verified prior to testing. All participants were administered segmental and suprasegmental speech perception tests, a frequency resolution task and receptive vocabulary assessment. This research was approved by the Human Research Protection Offices of Washington University School of Medicine in St. Louis and University of Miami in Miami, Florida.

#### **Subjects**

A total of 23 children (11 male; 12 female), meeting inclusion criteria, recruited from collaborators' patient populations (e.g., a pediatric audiology clinic), oral schools for the deaf and mainstream schools, were enrolled to participate in this research. Inclusion criteria

were: bilateral sensorineural HL, currently wear traditional HA(s) or previously wore traditional HA(s), use oral communication, are native to the English language and have normal cognition. Participants were between the chronological ages of 4 years 11 months and 11 years 11 months (mean=8.5 yrs; SD=2.4 yrs) at the time of testing. The recruiting clinician or facility reported normal cognition for all participants. Participant demographics are presented in Table 1 and the unaided better-ear hearing thresholds of the participants are presented in Figure 1. The children were identified with their HL at a range of ages (birth to 8 years of age; mean=2.4 yrs; SD=2.0 yrs) and a number of etiologies of HL were reported including: unknown (n=16), genetic causes (n=2), enlarged vestibular aqueduct (n=2), meningitis  $(n=1)$ , ototoxicity  $(n=1)$ , and cleft palate  $(n=1)$ . Nineteen of the participants wore bilateral HAs, 2 participants wore BiCROS devices, 1 participant only wore a right HA, and 1 participant with a slight to mild HL had discontinued HA use. Note that for the two participants using BiCROS HAs (P20 and P21), one has unaided thresholds (at the better ear) in the mild range in the low frequencies rising to within normal limits in the mid-to-high frequencies and the other participant has borderline normal/mild HL across frequencies at the better ear. Participant P13 has a steeply sloping HL with thresholds (at the better ear) within normal limits through 2000 Hz and a moderate-severe HL at 3000–8000 Hz. Participants utilized personal HAs from 3 manufacturers: Phonak (n=18), Oticon (n=3), and Widex (n=1), and were educated in a variety of settings, including: oral deaf education programs (n=13), mainstream school systems (n=9), and homeschool (n=1).

#### **Equipment and Procedures**

Testing was administered at Washington University School of Medicine in St. Louis, the Central Institute for the Deaf in St. Louis, and Hear Now in Miami, Florida. The Peabody Picture Vocabulary Test Version 4 (PPVT-4), the only standardized test in our battery, was administered according to the PPVT-4 Manual (Dunn and Dunn, 2007). All other test materials were computerized and administered with a Dell Latitude D630 or Panasonic PF-19 Toughbook computer. Auditory stimuli were delivered from the computer to an Anchor AN-100 speaker at approximately 60 dB SPL, as measured at the location of the listener's head, calibrated by a Quest Technologies 1200 Precision Integrating Sound Level Meter. Testing was completed in a single 2-to-2½ hour test session.

#### **Audibility Measures**

For each ear, unaided thresholds were obtained at octave frequencies from 250–8000 Hz and at 6000 Hz. Pure-tone averages (PTA: 500, 1000 and 2000 Hz) and low-frequency pure-tone averages (LF PTA; 250 and 500 Hz) were calculated at both ears for each participant. The participants' better-ear PTA and better-ear LF PTA were used for data analysis.

Electroacoustic analyses were conducted to verify the gain and output of the participants' HAs using Desired Sensation Level (DSL) targets (Scollie et al, 2005) and an Audioscan Verifit<sup>™</sup> system. SIIs (ANSI, 1997) were calculated by the Verifit system using the standard speech stimulus (carrot passage) for soft (50 dB SPL) and conversational/average (60 dB SPL) speech levels. All testing was completed using the participants' personal HAs, as programmed by their managing audiologist; no changes were made to the HAs prior to or following testing. The better-ear SIIs, computed for soft and average/conversational levels,

were used for data analyses. For the participant who had discontinued HA use, an unaided SII was calculated for the better ear.

#### **Outcome Measures**

Pediatric patients participated in tests which assessed segmental (Lexical Neighborhood Test [LNT] in quiet and noise) and suprasegmental (stress and talker discrimination) perception, frequency resolution ability (Spectral Modulation Detection [SMD]) and receptive vocabulary (Peabody Picture Vocabulary Test [PPVT]) abilities.

The LNT (Kirk et al, 1995) consists of two lists of 50 monosyllabic words (25 "lexically easy" and 25 "lexically hard") drawn from the vocabulary of 3-to-5-year-old typicallydeveloping children. List 1 was presented, in quiet, at a conversational level (60 dB SPL) and List 2 at a signal-to-noise ratio of +8 dB (speech at 60 dB SPL, four-talker babble at 52 dB SPL). The participant was instructed to repeat what was heard.

The within-female talker discrimination task assessed the ability to discriminate amongst female talkers based on voice characteristics (Geers et al, 2013). The recorded stimuli, a subset of the Indiana multi-talker speech database (Karl and Pisoni, 1994; Bradlow et al, 1996), included 8 adult females producing the Harvard IEEE sentences (IEEE Subcommittee, 1969). In each trial (32 trials total) the participant heard two different sentences and was instructed to indicate whether those sentences were spoken by the same female talker (16 trials) or different female talkers (16 trials).

The stress discrimination task is a newly-created test that assesses whether a child can discriminate stress patterns of bisyllabic (CVCV) nonwords. Recordings were made of an adult female producing two different phonetic bisyllabic nonwords ("badu" and "diga") with three different stress patterns (spondaic, trochaic, iambic). The digital recordings were modified slightly, using Antares AutoTune software, to create stimuli with the same total nonword duration, the same overall nonword intensity, and with specific acoustic properties to reflect syllabic stress. Compared to vowels in unstressed syllables, vowels in stressed syllables have elongated duration, increased intensity and increased fundamental frequency, F0. For spondaic stress nonwords, both vowels are about 250 ms in duration, have the same RMS, and approximately the same F0 of 145 Hz. For the trochaic stress nonwords, the first vowel is about 285 ms while the second is 142 ms in duration, the first vowel is about 5 dB more intense than the second, and the first vowel has a F0 about 40% higher than the second (202 Hz vs. 142 Hz). And, for the iambic stress nonwords, the first vowel is about 140 ms while the second is about 270 ms, the first vowel is about 5 dB less intense than the second, and the first vowel has a lower F0 than the second (148 Hz vs. 157 Hz). In each trial, two stimuli, with the same total duration, same RMS, and same phonetic sequence are presented. The participant was instructed to respond whether the two bisyllabic sequences have the "same" stress pattern or "different" stress patterns. In 30 total trials, 18 trials have the same stress pattern (6 trials each of spondee-spondee, trochee-trochee, iamb-iamb) and 12 trials have different stress patterns (4 trials each of trochee-spondee, iamb-spondee, trochee-iamb).

The SMD task assesses frequency resolution ability (Litvak et al, 2007; Saoji et al, 2009). Each trial is a 3-interval, 3-alternative forced-choice task. Three non-language, noise stimuli

are presented. One of the three stimuli in each trial is a noise with spectral modulation while the other two are noises with no spectral modulation (i.e., unmodulated spectrally, like white noise). Different spectral modulation frequencies and depths are mixed in the set of 60 trials employed here. All the noise stimuli were obtained from colleagues at Arizona State University (Spahr, personal communication) and are similar to those used by Eddins and Bero (2007) in a study with NH listeners. The stimuli were created by modifying a fouroctave wide (~350–5600 Hz) white noise with the application of a desired spectral modulation depth and frequency, and with a random starting phase in the spectral modulation. Inverse Fourier transformations of the noise spectrum produced a waveform with a specified spectral shape. Two spectral modulation frequencies, 0.5 and 1.0 cycles/ octave, and five spectral modulation depths,10, 11, 13, 14, and 16 dB, are represented equally in the set of 60 trials. All stimuli were 350 ms in duration and had the same overall level. The participant was instructed to choose the noise that sounded different from the other two.

The PPVT, fourth edition, assesses receptive vocabulary ability (Dunn and Dunn, 2007). Words of increasing difficulty were presented verbally along with a page of four pictures. Participants were instructed to point to the picture best representing the word presented.

#### **Statistical Analysis**

Pearson correlations were computed to examine the relations between measures of audibility (LF PTA, PTA, and SII) and speech perception (suprasegmental and segmental); and frequency resolution ability (SMD) and speech perception. Correlations between SMD and receptive vocabulary were also examined. The correlations between age and measures of audibility, frequency resolution and speech perception were examined. More elaborate statistical analyses were not deemed appropriate due to a relatively small sample size and to violations of linear regression assumptions (e.g., non-normal distributions for some of the data and collinearity among some variables).

## **RESULTS**

As shown in the audiogram in Figure 1, participants had a wide range of HL. The average better-ear PTA was 45.4 dB HL (SD=18.3) and the average better-ear LF PTA was 36.0 dB HL (SD=18.2). These values reflect the general trend for children to have better thresholds in the low to mid frequency range compared to thresholds at high frequencies. Examination of Verifit fitting data revealed that of the 22 children utilizing hearing aids in this study, all but two children had outputs at soft and conversation levels speech that met DSL targets at 250–4000 Hz for at least one ear. Figure 2 shows the distributions of SIIs for soft and conversational speech. The mean SII for soft speech (50 dB SPL) was 65.0 (SD=17.9) while the mean SII for conversational speech (60 dB SPL) was 75.4 (SD=14.5).

The scores for the LNT words in quiet and noise are shown in the box plots in Figure 3. The mean LNT word score in quiet was 84% correct (SD=16) and in noise, 69% correct  $(SD=16)$ .

For the suprasegmental perception tests (see Figure 3) most participants scored above chance levels of performance (50% correct for both talker and stress discrimination). The mean score for the stress discrimination task was 82% correct (SD=17) with 22 participants scoring above chance. Twenty participants scored above chance on the talker discrimination task with a mean score of 68% correct (SD=16). Twenty participants scored above chance for the SMD task (33% correct), with a mean score of 61% correct (SD=24).

Raw and standard scores were computed for the receptive vocabulary (PPVT) results. A raw score corresponds to the number of items below a basal level plus the number of items correctly selected. This raw score can be converted to a standard score relative to the ageappropriate normative group, where the mean is set at 100 with a standard deviation of 15. Both the raw and standard scores of the PPVT are reported and used for analysis in this research because all other test measures are not standardized. The standard scores of the participants in this study ranged from 75–124 (mean= 94.2; SD= 13.1). Five of the participants scored at or above 100 (the normative mean for NH children), 13 of the participants scored within one standard deviation below the normative mean (i.e., had standard scores in the 85–100 range) and 5 scored within one-to-two standard deviations below the normative sample (i.e., had standard scores in the 70–85 range).

Correlation values amongst demographic characteristics (age at test), measures of audibility (PTAs, SIIs), perception scores (segmental, suprasegmental, frequency-resolution ability), and receptive language ability are presented in the appendix. All outcome measures (LNT in quiet, LNT in noise, stress discrimination, talker discrimination, SMD, PPVT raw scores) were significantly correlated with age (r values range from 0.41 to 0.85, p values from  $\lt$ 0.0001 to 0.05). As the children increased in age, their performance on these tasks improved. None of the measures of audibility (PTA, LF PTA, SII-soft, SII-avg) were significantly correlated with age. As expected, the audiological variables (PTA, SII-soft, SII-avg) were significantly correlated with each other (r values range, in magnitude, from 0.46 to 0.94; p values range from < 0.0001 to 0.028). As might also be expected, the measures of segmental perception were significantly correlated with each other (LNT in quiet and LNT in noise;  $r =$ 0.84,  $p < 0.001$ ) as were the measures of suprasegmental perception (stress discrimination and talker discrimination;  $r = 0.80$ ,  $p < 0.001$ ).

LF PTA was significantly correlated with suprasegmental perception, both stress discrimination ( $r = -0.50$ ,  $p < 0.01$ ) and talker discrimination ( $r = -0.47$ ,  $p < 0.05$ ). The scatterplot in Figure 4 shows LF PTA versus talker discrimination performance: participants with better LF PTAs have better talker discrimination scores. Traditional PTA was also significantly correlated with stress discrimination ( $r = -0.48$ ,  $p < 0.05$ ), but was not significantly correlated with talker discrimination. However, the measures of audibility (PTA, LF PTA, SII-soft, SII-avg) and segmental speech perception (LNT in quiet and LNT in noise) were not significantly correlated. The scatterplot in Figure 5 shows an example relation that is not significantly correlated, namely LF PTA versus performance on the LNT in noise.

The SMD scores were significantly correlated with both PTA ( $r = -0.57$ ,  $p < 0.01$ ) and LF PTA ( $r = -0.59$ ,  $p < 0.01$ ), but were not significantly correlated with SIIs. SMD scores were

also significantly correlated with measures of both suprasegmental and segmental speech perception as well as with receptive language scores, both raw and standard (r values range from 0.49 to 0.77, p values range from  $< 0.0001$  to 0.018). The scatterplot in Figure 6 shows performance on the SMD task versus PPVT standard score (score relative to the ageappropriate normative group). Children who performed better on the SMD task had better standardized receptive vocabulary scores.

## **DISCUSSION**

Clinicians tend to focus on the relationship between audibility and segmental speech perception (perception of "what is said"). The relationship between audibility and suprasegmental perception (perception of "who said it" and "how it is said") has been lessstudied and is not as well understood. Perception of suprasegmental features of speech is undeniably important though, in particular for early speech and language development (Werker and Yeung, 2005), as well as for later conversational benefits of recognizing who the speaker is and the intent of the spoken message. Further understanding this relationship might guide clinicians' recommendations for hearing devices for children with HL as they develop spoken language. In addition, measures of frequency resolution have been thought to assist in predicting benefit from amplification (Bernstein et al, 2013). Learning more about frequency resolution ability, assessed with non-linguistic stimuli, and its relationship to speech perception might assist audiologists in making hearing device recommendations (Rakita, 2012; Bernstein et al, 2013). The purpose of this study was to explore these relationships between audibility, and segmental and suprasegmental speech perception, as well as the relationship between audibility, frequency resolution ability and speech perception in children who use HAs. These results may be particularly important when considering moving from bilateral HAs to bimodal devices, as the HA is an important component of the bimodal device configuration, providing good LF information for bimodal benefit (Nittrouer and Chapman, 2009; Zhang et al, 2013).

In contrast to the results of other studies (Davis et al, 1986; Delage and Tuller, 2007; Stiles et al, 2012), including the work of Davidson and Skinner (2006), this current study did not find a relationship between audibility (unaided PTA or aided SII) and segmental perception. The absence of this relationship in this current research may be attributed to a number of factors. Most notable is the high level of overall speech recognition performance, particularly for the LNT words presented in quiet. The mean score for the HA group is 84% correct, with only 3 participants' scores less than 70%. Even when the LNT words are presented in noise, recognition remains relatively high (69% correct, with only 4 participants' scores less than 60%). The participants in the Davidson and Skinner study (2006), in general, had less residual hearing and poorer scores on the LNT in quiet (mean 54%, range 10–85%). Furthermore, aided SIIs for conversational speech (60 dB SPL) were generally high (SII range: 34–98). And, it is important to note that the majority of the participants' (n=19,  $\sim$ 83%) aided SIIs for conversational speech were  $\sim$  65, the minimum aided SII that has been suggested for developing spoken language skills of children with HL (recent pediatric studies; Stiles et al, 2012, McCreery et al, 2013). Thus, the relatively high levels of word recognition in combination with relatively well-fit HAs and, in turn, the

limited variability in aided audibility may have precluded finding a relationship between audibility and segmental speech perception.

While for these participants there were no significant relationships between audibility measures and segmental perception, there were relationships of interest between audibility measures and suprasegmental perception. In particular, LF PTA was significantly correlated with suprasegmental perception. Suprasegmental features of speech are primarily carried in voice-pitch related properties which are most prominent at low frequencies. This result further emphasizes the importance of the low-frequency audibility that HAs can provide, as demonstrated in numerous previous studies (Tyler et al, 2002; Carroll and Zeng, 2007; Ching et al, 2007; Most and Peled, 2007; Dorman et al, 2008; Nittrouer and Chapman, 2009; Most et al, 2011; Sheffield and Zeng, 2012). Access to low-frequency sound energy via a HA in-turn gives access to suprasegmental features of speech (i.e., the acoustic cues of intonation, stress and rhythm), which may allow the HA user to attend to the same cues that NH children use in their early language development (Jusczyk et al, 1999; Werker and Yeung, 2005; Seidl and Cristia, 2008; Seidl and Johnson, 2008; Swingley, 2009). The perception of these suprasegmental features is not only the basis of early language processing and vocabulary acquisition, but is also needed for successful communication throughout life. For example, good suprasegmental perception gives a listener the ability to discern who is speaking, and the ability to identify any overlaid emotional and/or sarcastic intent. Suprasegmental speech perception tests may become more relevant as children with greater degrees of residual hearing are considered for cochlear implants. Many of these children will continue to use a HA at the non-implanted ear, especially those with good lowfrequency residual hearing thresholds (Firszt et al, 2012; Cadieux et al, 2013; Davidson et al, 2015).

Results from the SMD task were of particular interest. Performing well on this task requires good supra-threshold hearing abilities; abilities beyond detection, the gold standard of audiometric testing (the traditional audiogram). Participants did not simply signal that they heard the three sounds (detection) but also that they could hear a difference amongst the three sounds presented (one with spectral modulation and two without). Hence, this task requires a discrimination ability, in this case a frequency-resolution ability (due to the wideband, spectrally modulated - or not - nature of the stimuli). It has been suggested previously that performance on this task may be useful in predicting bimodal benefit (Bernstein et al, 2013; Zhang et al, 2013). Others have found a relationship between SMD performance and segmental perception (Saoji et al, 2009; Spahr et al, 2011; Zhang et al, 2013). In this current study, SMD was significantly correlated with the threshold measures of audibility (LF PTA and PTA) and with most outcome measures (PPVT raw and standard scores, LNT in noise, and stress and talker discrimination). There was also a trend toward a significant correlation with LNT in quiet, a relationship that might have been seen with more participants (n=23). These results indicate that performance on the SMD task is not only related to segmental perception, but also has a strong relation with suprasegmental perception.

Although the presumed non-linguistic nature of the SMD task makes it a highly desirable test for children with HL, its potential usefulness may be tempered somewhat by the

presence of a very strong relationship between SMD performance and age at test. This relationship indicates a developmental effect. This effect has also been reported by Rakita (2012) and Bridges (2014) for children with both normal and impaired hearing of similar ages. That is, as age increases, SMD performance (a type of spectral resolution ability) also increases. However, it is unclear whether this relationship is due to a developmental improvement in frequency resolving ability (maturation in the auditory system) or to a developmental improvement in the ability to perform behavioral tasks, including psychoacoustic ones (due to a maturation in language skills and the ability to understand a task), or both. This task required concentration for about 10–20 minutes and also required periods of pausing for re-instruction for the younger participants. These could be important factors to consider if using this task clinically.

The relationship between SMD and receptive vocabulary (PPVT) standard scores requires further investigation as to its utility and interpretation with young pediatric patients. SMD performance was significantly correlated with both the raw and standard scores of the PPVT. Since there is an apparent developmental age effect with the SMD task, the relationship between SMD performance and PPVT raw scores was not surprising. However, the PPVT standard scores take age into account (i.e., the developmental age effect is not carrying the significant correlation). Although smaller than that with PPVT raw scores ( $r = 0.59$  vs  $r =$ 0.75), the significant relationship between SMD and the PPVT standard score suggests other interpretations. Although correlation does not specify causality, it might be that good frequency resolution (as measured by SMD performance) leads to good speech perception ability, which then leads to good vocabulary. These relations warrant further investigation, with larger numbers of participants of various ages. A time-efficient, non-linguistic measure that predicts device benefit would be invaluable to clinicians, especially those who work with children (who can be difficult-to-test and may have limited language).

In addition to observing an age effect with performance on the SMD task, relationships with age were seen with all other outcome measures as well (LNT in quiet and noise, stress and talker discrimination). That is, as the age of the child increased, performance also increased. In this study, the children ranged in age from approximately 5 to 11 years at the time of testing. Again, the relationship between age and other perceptual outcome measures may be due to auditory maturation or maturation in the ability to concentrate and/or understand the task, or all of these.

## **CONCLUSIONS**

These results may be important for helping clinicians decide how much hearing is needed for and potential benefit with selected hearing devices. This will be especially true for children with more severe HL who may be candidates for one or two CI(s), in order to recommend the appropriate hearing devices for maximizing speech and language development. HAs provide low frequency information that is key for perceiving suprasegmental information, and a period with HA(s) prior to cochlear implantation may be beneficial.

This current research supports the idea that HAs provide important low-frequency information, which facilitates suprasegmental speech perception. However, these results must be considered with caution as there were only 23 participants enrolled in this study. Further research is needed to clarify the underlying mechanisms and developmental effects found with the SMD task. The potential clinical use of SMD and suprasegmental tasks is great, and further investigations using these types of tests with children with HL are warranted.

## **Acknowledgments**

**Funding Sources:** This publication was supported by the Washington University Institute of Clinical and Translational Sciences grants UL1 TR000448 and TL1 TR000449 from the National Center for Advancing Translational Sciences and grant R01 DC012778 (Davidson-PI) from the National Institute on Deafness and Other Communication Disorders.

We are grateful to all the children and their families for graciously giving their time to participate in this study. We also thank Chris Brenner and Sarah Fessenden for all their contributions to this project, especially their assistance in subject recruitment and data analysis.

## **ABBREVIATIONS**



### **References**

Abercrombie, D. Elements of general phonetics. Chicago: Aldine Pub; 1967.

- American Academy of Audiology. American Academy of Audology Clinical Practice Guidelines: Pediatric Amplification. Reston, VA: American Academy of Audiology; 2013.
- ANSI. ANSI S3.5–1997, American National Standard Methods for Calculation of the Speech Intelligibility Index. New York, NY: American National Standards Institute; 1997.
- Bernstein JGW, Mehraei G, Shamma S, Gallun FJ, Theodoroff SM, Leek MR. Spectrotemporal modulation sensitivity as a predictor of speech intelligibility for hearing-impaired listeners. J Am Acad Audiol. 2013; 24(4):293–306. [PubMed: 23636210]
- Blamey PJ, Sarant JZ, Paatsch LE, Barry JG, Bow CP, Wales RJ, Wright M, Psarros C, Rattigan K, Tooher R. Relationships among speech perception, production, language, hearing loss, and age in children with impaired hearing. J Speech Lang Hear Res. 2001; 44(2):264–285. [PubMed: 11324650]
- Boothroyd A. Auditory perception of speech contrasts by subjects with sensorineural hearing loss. J Speech Lang Hear Res. 1984; 27:134–144.
- Boothroyd A, Eran O. Auditory speech perception capacity of child implant users expressed as equivalent hearing loss. Volta Review. 1994; 96(5):17.
- Bradlow AR, Torretta GM, Pisoni DB. Intelligibility of normal speech I: Global and fine-grained acoustic-phonetic talker characteristics. Speech Commun. 1996; 20(3):255–272. [PubMed: 21461127]
- Bridges, HJ. Paper 697. Program in Audiology and Communication Sciences. Washington University School of Medicine; 2014. Spectral resolution and speech understanding in children and young adults with bimodal devices. Independent Studies and Capstones.
- Cadieux JH, Firszt JB, Reeder RM. Cochlear implantation in nontraditional candidates: preliminary results in adolescents with asymmetric hearing loss. Otol Neurotol. 2013; 34(3):408–415. [PubMed: 23222962]
- Carroll J, Zeng FG. Fundamental frequency discrimination and speech perception in noise in cochlear implant simulations. Hear Res. 2007; 231(1–2):42–53. [PubMed: 17604581]
- van Wanrooy, Ching E., Dillon, H. Binaural-bimodal fitting or bilateral implantation for managing severe to profound deafness: a review. Trends Amplif. 2007; 11(3):161–192. [PubMed: 17709573]
- Davidson LS, Skinner MW. Audibility and speech perception of children using wide dynamic range compression hearing aids. Am J Audiol. 2006; 15(2):141–153. [PubMed: 17182879]
- Davidson LS, Firszt JB, Brenner C, Cadieux JH. Evaluation of hearing aid frequency response fittings in pediatric and young adult bimodal recipients. J Am Acad Audiol. 2015; 26(4):393–407. [PubMed: 25879243]
- Davis JM, Elfenbein J, Schum R, Bentler RA. Effects of mild and moderate hearing impairments on language, educational, and psychosocial behavior of children. J Speech Hear Disord. 1986; 51:53– 62. [PubMed: 3945060]
- Delage H, Tuller L. Language development and mild-to-moderate hearing loss: Does language normalize with age? J of Speech Lang and Hear Res. 2007; 50:1300–1313. [PubMed: 17905913]
- Dorman MF, Gifford RH, Spahr AJ, McKarns SA. The benefits of combining acoustic and electric stimulation for the recognition of speech, voice and melodies. Audiol Neurootol. 2008; 13(2):105. [PubMed: 18057874]
- Dorman MF, Cook S, Spahr A, Zhang T, Loiselle L, Schramm D, Whittingham J, Gifford R. Factors constraining the benefit to speech understanding of combining information from low-frequency hearing and a cochlear implant. Hear Res. 2015; 322:107–111. [PubMed: 25285624]
- Dunn CC, Walker EA, Oleson J, Kentworthy M, Van Voorst T, Tomblin JB, Ji H, Kirk KI, McMurray B, Hanson M, Gantz BJ. Longitudinal speech perception and language performance in pediatric cochlear implant users: The effect of age at implantation. Ear Hear. 2014; 35(2):148–160. [PubMed: 24231628]
- Dunn LM, Dunn DM. 2007PPVT-4: Peabody picture vocabulary test.
- Eddins DA, Bero EM. Spectral modulation detection as a function of modulation frequency, carrier bandwidth, and carrier frequency region. J Acoust Soc Am. 2007; 121(1):363–372. [PubMed: 17297791]
- Erber NP. Auditory, visual, and auditory-visual recognition of consonants by children with normal and impaired hearing. J Speech Hear Res. 1972a; 15(2):413–422. [PubMed: 5047880]

- Erber NP. Speech-envelope cues as an acoustic aid to lipreading for profoundly deaf children. J Acoust Soc Am. 1972b; 51(4):1224–1227. [PubMed: 4113475]
- Erber NP. Speech Perception by Profoundly Hearing-Impaired Children. J Speech Hear Disord. 1979; 44(3):255–270. [PubMed: 480932]
- Eisenberg LS, Martinez AS, Sennarogly G, Osberger MJ. Establishing new criteria in selecting children for a cochlear implant: Performance of "platinum" hearing aid users. Ann Otol Rhinol Laryngol Suppl. 2000; 185:30–33. [PubMed: 11140994]
- Firszt JB, Holden LK, Reeder RM, Cowdrey L, King S. Cochlear implantation in adults with asymmetric hearing loss. Ear Hear. 2012; 33(4):521–533. [PubMed: 22441359]
- Geers AE, Davidson LS, Uchanski RM, Nicholas JG. Interdependence of linguistic and indexical speech perception skills in school-age children with early cochlear implantation. Ear Hear. 2013; 34(5):562–574. [PubMed: 23652814]
- Grant KW. Identification of intonation contours by normally hearing and profoundly hearing-impaired listeners. J Acoust Soc Am. 1987; 82(4):1172–1178. [PubMed: 3680776]
- Hack ZC, Erber NP. Auditory, visual, and auditory-visual perception of vowels by hearing-impaired children. J Speech Hear Res. 1982; 25(1):100–107. [PubMed: 7087410]
- IEEE Subcommittee. IEEE Recommended Practice for Speech Quality Measurements. IEEE Trans. on Audio and Electroacoustics. 1969:225–246.
- Illg A, Bojanowicz M, Lesinski-Schiedat A, Lenarz T, Buchner A. Evaluation of the bimodal benefit in a large cohort of cochlear implant subjects using a contralateral hearing aid. Otol Neurotol. 2014; 35(9):240–244.
- Jusczyk PW, Houston DM, Newsome M. The beginnings of word segmentation in English-learning infants. Cogn Psychol. 1999; 39(3–4):159–207. [PubMed: 10631011]
- Karl JR, Pisoni DB. Effects of stimulus variability on recall of spoken sentences: A first report. Research on Spoken Language Processing Progress Report No. 1994; 19:145–193.
- Kirk KI, Pisoni DB, Osberger MJ. Lexical effects on spoken word recognition by pediatric cochlear implant users. Ear Hear. 1995; 16:470–481. [PubMed: 8654902]
- Kong YY, Stickney GS, Zeng FG. Speech and melody recognition in binaurally combined acoustic and electric hearing. J Acoust Soc Am. 2005; 117(3):1351–1361. [PubMed: 15807023]
- Lederberg AR, Prezbindowski AK, Spencer PE. Word-learning skills of deaf preschoolers: the development of novel mapping and rapid word-learning strategies. J Child Dev. 2000; 71(6):1571– 1585.
- Litvak L, Spahr A, Saoji A. Relationship between perception of spectral ripple and speech recognition in cochlear implant and vocoder listeners. J Acoust Soc Am. 2007; 122(2):982–991. [PubMed: 17672646]
- Massicotte-Laforge S, Shi R. The role of prosody in infants' early syntactic analysis and grammatical categorization. J Acoust Soc Am. 2015; 138(4):EL441–EL446. [PubMed: 26520358]
- McCreery RW, Bentler RA, Roush PA. Characteristics of hearing aid fittings in infants and young children. Ear Hear. 2013; 34(6):701–710. [PubMed: 23575463]
- Most T, Harel T, Shpak T, Luntz T. Perception of suprasegmental speech features via bimodal stimulation: Cochlear implant on one ear and hearing aid on the other. J Speech Lang Hear Res. 2011; 54:668–678. [PubMed: 20844254]
- Most T, Peled M. Perception of suprasegmental features of speech by children with cochlear implants and children with hearing aids. J Deaf Stud Deaf Educ. 2007; 12(3):350–361. [PubMed: 17478417]
- Nittrouer S, Chapman C. The effects of bilateral electric and bimodal electric-acoustic stimulation on language development. Trends Amplif. 2009; 13(3):190–205. [PubMed: 19713210]
- Osberger MJ, Miyamoto RT, Zimmerman-Phillips S, Kemink JL, Stroer BS, Firszt JB, Novak MA. Independent evaluation of the speech perception abilities of children with the Nucleus 22-channel cochlear implant system. Ear Hear. 1991; 12(4 Suppl):66S–80S. [PubMed: 1955092]
- Pisoni, DB. Some thoughts on "normalization" in speech perception. In: Johnson, K., Mullennix, JW., editors. Talker variability in speech processing. San Diego, CA: Academic Press; 1997.

- Rakita, L. Paper 655. Program in Audiology and Communication Sciences. Washington University School of Medicine; 2012. Spectral modulation detection in normal hearing children. Independent Studies and Capstones.
- Saoji A, Litvak L, Spahr AJ, Eddins DA. Spectral modulation detection and vowel and consonant identifications in cochlear implant listeners. J Acoust Soc Am. 2009; 126(3):955–958. [PubMed: 19739707]
- Scollie SD. Children's speech recognition scores: the Speech Intelligibility Index and proficiency factors for age and hearing level. Ear Hear. 2008; 29(4):543–556. [PubMed: 18469717]
- Scollie S, Seewald R, Cornelisse L, Moodie S, Bagatto M, Laurnagaray D, Beaulac S, Punford J. The desired sensation level multistage input/output algorithm. Trends Amplif. 2005; 9(4):159–197. [PubMed: 16424945]
- Seidl A, Cristia A. Developmental changes in the weighting of prosodic cues. Dev Sci. 2008; 11(4): 596–606. [PubMed: 18576967]
- Seidl A, Johnson EK. Boundary alignment enables 11-month-olds to segment vowel initial words from speech. J Child Lang. 2008; 35(1):1–24. [PubMed: 18300427]
- Shannon RV, Zeng FG, Kamath V, Wygonski J, Ekelid M. Speech recognition with primarily temporal cues. Science. 1995; 270:303–304. [PubMed: 7569981]
- Sheffield BM, Zeng FG. The relative phonetic contributions of a cochlear implant and residual acoustic hearing to bimodal speech perception). J Acoust Soc Am. 2012; 131(1):518–530. [PubMed: 22280613]
- Spahr A, Saoji A, Litvak L, Dorman M. Spectral cues for understanding speech in quiet and in noise. Cochlear Implants Int. 2011; 12(Suppl 1):S66–S69. [PubMed: 21756477]
- Stiles DJ, Bentler RA, McGregor KK. The Speech Intelligibility Index and the pure-tone average as predictors of lexical ability in children fit with hearing aids. J Speech Lang Hear Res. 2012; 55(3): 764–778. [PubMed: 22223888]
- Swingley D. Contributions of infant word learning to language development. Philos Trans R Soc Lond B Biol Sci. 2009; 364(1536):3617–3632. [PubMed: 19933136]
- Thiessen ED, Hill EA, Saffran JR. Infant-directed speech facilitates word segmentation. Infancy. 2005; 7(1):53–71.
- Tyler RS, Parkinson AJ, Wilson BS, Witt S, Preece JP, Noble W. Patients utilizing a hearing aid and a cochlear implant: speech perception and localization. Ear Hear. 2002; 23(2):98–105. [PubMed: 11951854]
- Van Tasell DJ, Greenfield DG, Logemann JJ, Nelson DA. Temporal cues for consonant recognition: Training talker generalization, and use in evaluation of cochlear implants. J Acoust Soc Am. 1992; 92(3):1247–1257. [PubMed: 1401513]
- Walker EA, Holte L, McCreery RW, Spratford M, Page T, Moeller MP. The influence of hearing aid use on outcomes of children with mild hearing loss. J Speech Lang Hear Res. 2015; 58(5):1611– 1625. [PubMed: 26151927]
- Werker JF, Yeung HH. Infant speech perception bootstraps word learning. Trends Cogn Sci. 2005; 9(11):519–527. [PubMed: 16202639]
- Wolfe J, John A, Schafer E, Hudson M, Boretzki M, Scollie S, Woods W, Wheeler J, Hudgens K, Neumann S. Evaluation of wideband frequency responses and non-linear frequency compression for children with mild to moderate high-frequency hearing loss. Int J Audiol. 2015; 54:170–181. [PubMed: 25731582]
- Won J, Drennan W, Rubinstein J. Spectral-Ripple Resolution Correlates with Speech Reception in Noise in Cochlear Implant Users. J Assoc Res Otolaryngol. 2007; 8:384–392. [PubMed: 17587137]
- Zhang T, Spahr AJ, Dorman MF, Saoji A. Relationship between auditory function of nonimplanted ears and bimodal benefit. Ear Hear. 2013; 34(2):133–141. [PubMed: 23075632]





Individual audiometric better-ear hearing thresholds. The better-ear unaided hearing thresholds, 250–8000 Hz, of each participant are plotted on the audiogram.



### **Figure 2.**

Box-plots for Speech Intelligibility Indices at soft and average inputs for the 23 participants' better ears. The limits of the box represent the lower and upper quartile of the distribution (the difference is the interquartile range, IQR), and the horizontal line through the box represents the median. Whiskers represent the minimum and maximum scores in the distribution, excluding any outliers. Open-circle outliers are values between 1.5 and 3 IQRs from the end of a box; asterisk outliers are values more than 3 IQRs from the end of a box.



#### **Figure 3.**

Box-plots for the segmental and suprasegmental speech perception measures, LNT Words in Quiet, LNT Words in Noise, Talker Discrimination, and Stress Discrimination. Details about box-plots are the same as those in Figure 2.



#### **Figure 4.**

LF PTA hearing thresholds (dB HL) vs performance on the talker discrimination task. Individual data points are represented by open circles.





LF PTA hearing thresholds vs. performance on the stress discrimination task. Individual data points are represented by open circles.



#### **Figure 6.**

Performance on the SMD task vs. PPVT standard score. Individual data points are represented by open circles.

**Table 1**

Participant demographics Participant demographics

