2011

The effect of talker age and gender on speech perception of pediatric hearing aid users

Elizabeth J. Robinson

THE EFFECT OF TALKER AGE AND GENDER ON SPEECH PERCEPTION OF PEDIATRIC HEARING AID USERS

by

Elizabeth J. Robinson

A Capstone Project
submitted in partial fulfillment
of the requirements for the degree of:

Doctor of Audiology

Washington University School of Medicine
Program in Audiology and Communication Sciences

May 18, 2012

Approved by:
Lisa S. Davidson, Ph.D., Capstone Project Advisor
Rosalie M. Uchanski, Ph.D., Secondary Reader

***Abstract:** Even though pediatric hearing aid (HA) users listen most often to female talkers, clinically-used speech tests primarily consist of adult male talkers' speech. Potential effects of age and/or gender of the talker on speech perception of pediatric HA users were examined using two speech tests, hVd-vowel identification and CNC word recognition, and using speech materials spoken by four talker types (adult males, adult females, 10-12 year old girls, and 5-7 year old girls). For the nine pediatric HA users tested, word scores for the male talker's speech were higher than those for the female talkers, indicating that talker type can affect word recognition scores and that clinical tests may over-estimate everyday speech communication abilities of pediatric HA users.*

Acknowledgements

Partial support provided by NIH NIDCD Grant K23 DC008294. The author acknowledges: Lisa S. Davidson, Ph.D. and Rosalie M. Uchanski, Ph.D. for their constant guidance, time and endless patience throughout this past year; Christine Brenner for her continual support and help in data management; Margaret Pikora and the audiologists at Special School District as well as Marie Richter and Susie Whaley at the Central Institute for the Deaf for their help in the recruitment process; Dorina Kallogjeri, M.P.H. for assistance with statistical analysis; Peter Assmann, Ph.D. and William Katz, Ph.D. at University of Texas- Dallas for use of their hVd stimuli; James Hillenbrand, Ph.D. at Western Michigan University for use of his hVd stimuli; Maddie Fessenden for recording of the CNC stimuli; and finally, all of the subjects and their parents for the time and participation in this study

Table of Contents

List of Tables & Figures

Introduction

To quantify auditory performance or improvement over time by hearing-impaired (HI) individuals, clinical audiologists commonly utilize speech perception tests. For HI individuals who use amplification devices such as cochlear implants (CI) and hearing aids (HA), results from these tests may direct programming adjustments to the device in an attempt to maximize an individual's access to speech signals. It is assumed, although often not stated explicitly, that results from these clinical speech perception tests will represent the HI individual's communication abilities in everyday situations. However, for young children this may not be the case. Most young children spend the majority of their day listening to adult female talkers, to other children, and to themselves (self-monitoring), but not to the adult male speech of commonly-used speech perception tests.

Clinical speech perception results may vary greatly due to a combination of factors related to the patient, the test and the tester (Boothroyd, 2004). Factors specific to the patient include cause and onset of hearing loss (pre-lingual or post-lingual), language level of patient, communication mode (oral or sign), and amplification devices (CI, HA, CI and HA combined). Factors related to the test include type of speech materials (nonsense syllables, words, or sentences), type of presentation (live-voice or recorded), predictability of speech materials (high-context or low-context), and testing environment (noise or quiet). The format of the test can also influence results. For example, open-set tests are more demanding and require higher cognitive processing abilities than closed-set tests (Eisenberg et al., 2005). In addition, another test-related factor is the type of talker (e.g., talker's age and gender) who produces the speech used in the speech tests.

Speech signals vary acoustically across gender and age (Eguchi & Hirsh, 1969; Lee et al., 1999). Anatomical and physiological changes during development are largely responsible for many of these acoustic differences. During typical development, the larynx descends, the vocal folds increase in size and mass, and the vocal tract lengthens. The average fundamental frequencies (f0) for adult males, adult females, and children (age 3) are approximately 124, 221, and 300 Hz respectively, due mainly to the size and mass of the vocal folds. It is known that as children develop and mature, f0 gradually decreases. The average f0 for both 7- year-old boys and girls is 262 Hz, for 13-year-old girls is 240 Hz, and for 13-year-old boys is 221 Hz, while for adult females and males, they are 221 Hz and 124 Hz, respectively (Eguchi & Hirsh, 1969). Hillenbrand et al. (1995) measured acoustic characteristics of vowels for a variety of talkers, and report an average f0 for adult males of 130 Hz, for adult females, 220 Hz, and for 10- to 12-year old boys and girls, 236 Hz. Across studies, the data show consistently that f0 is highest at the youngest ages, and is similar for boys and girls until adolescence.

Specific speech sounds also have different acoustic characteristics, depending on talker type. The average second formant (F2) frequency values for the vowel /i/ (which has the highest F2 of all English vowels) are approximately 2217, 2810, and 3108 Hz, for adult males, adult females and children (Eguchi & Hirsh, 1969) and 3400-3535 Hz for 3, 5, and 7- year old children (Assmann & Katz, 2000). For any individual talker, the locations of the spectral peaks of the fricative sounds /s/ and /sh/ differ due to their different places of articulation in the vocal tract. The frequencies of these spectral peaks are dependent on the overall length of the vocal tract, which, varies for children vs. adult talkers, and female vs. male talkers. In Jongman et al. (2000), the spectral peaks of /s/ and /sh/ are reported as 6.2 and 3.3 kHz, for adult males, and 7.5 and 4.3 kHz for adult females. Using a 1/3-octave band analysis, the spectral peaks of /s/ and

/sh/ are reported as 5.4 and 3 kHz for adult males, and 7.3 and 4.5 kHz for adult female talkers (Pittman et al., 2003). Also, for 2-4 year old children, Pittman et al. (2003) report that the spectral peaks for /sh/ are about 4.5 kHz, and that peaks for /s/ are beyond 8 kHz, or the limits of their 1/3-octave-band analysis. Additionally, long-term average speech spectra (LTASS) differ for different types of talkers. For speakers of American English, there is more energy in the mid-frequencies (1-2 kHz) for adult males than for females, though there is some variation with geographic location (Byrne et al., 1994). Compared to children, Pittman et al. (2003) found the LTASS to be 7-9 dB higher for frequencies above 2 kHz for the adult talkers. Speech produced by adults and children exhibit other differences, including increased variability in articulation timing by children in comparison to adult talkers, most likely due to a lack of maturity of the speech motor control in children (Eguchi & Hirsh, 1969; Lee et al., 1999).

Pearson et al. (1977) measured teachers' speech levels and background noise levels in classrooms providing signal-to-noise ratios of "everyday" listening situations. Speech spectra of females, males, and children were measured in a quiet environment at different vocal efforts; casual, normal, raised, loud and shouted speech. The results showed the speech spectrum of male speakers have the most energy at 125 Hz, but that in the frequency range 250 – 8000 Hz there is not much difference amongst the three speakers for casual, normal and raised speech. For males, loud and shouted speech have increased energy in the 1250 – 1600 Hz region, while female and children's voices were very similar across all speaking styles.

Speech is usually considered the most important class of sounds for all people, but is especially so for children during their critical period of speech and language acquisition. For HI individuals, cochlear implants (CI) and hearing aids (HA) should provide access to all acoustic signals. To assess benefit provided by amplification, audiologists use speech perception tests to record progress of auditory development. Adequate bandwidth (BW), typically determined by the upper end of the frequency response of a hearing aid, is an integral factor in the audibility of sounds with predominately high-frequency energy, such as the fricative sounds /s/ and /z/. Commonly, the wider the bandwidth, the higher the HA user will score on speech perception tests in quiet and in noise (Skinner & Miller, 1983). Because the two phonemes, /s/ and /z/, in English, mark plurality, possession and verb tense, they play an important role in a child's development of spoken language.

Results from previous studies indicate that children and adults may have different requirements regarding BW for good speech perception. Kortekaas & Stelmachowicz (2000) low-pass filtered the phoneme /s/ and presented it to normal hearing adults and children. Results showed that children required a wider bandwidth than adults to correctly perceive the high frequency phoneme /s/. Stelmachowicz et al. (2001) explored the effect of bandwidth on the perception of /s/ by normal hearing (NH) and hearing impaired (HI) children and adults. Stimuli were consonant-vowel (CV) and vowel-consonant (VC) syllables spoken by an adult male, an adult female, and a child. Closed-set responses to these syllables were gathered. Stimuli were low-pass filtered (LPF) at six different cutoff frequencies (2, 3, 4, 5, 6, and 9 kHz) to create six listening conditions. For the adult male talker, NH children and NH adults performed better than 80% correct at a BW of 4 kHz. However, HI children and HI adults required a BW of 5 kHz to reach an 80% correct performance level. For the female talker; performance across all groups was 33% correct for the 2, 3, 4, and 5 kHz BWs, and then increased drastically at 6 kHz for both NH groups and at 9 kHz for the HI groups. At a BW of 9 kHz, the condition that yielded the highest scores, the NH groups and HI adults scored better than 80% correct while the HI children scored 77% correct. For the child talker, steady improvements were seen across all listener groups as BW increased. NH groups and HI adults achieved scores greater than 90% correct at 9 kHz while the HI children only reached 75% correct for female and child talkers and 80% for male talkers. In each condition, for all talkers and across listener groups, scores increased as BW increased. These perceptual results can be explained given the spectral characteristics of these three talkers' /s/ sounds. The peaks in spectral energy of their /s/ sounds occur at 4.2 - 6.9 kHz for the adult male, and at 6.3 – 8.8 kHz for the adult female and child.

Stelmachowicz et al. (2002) assessed how well children with HI, wearing their personal hearing aid/s, perceive the fricatives /s/ and /z/ in the plural form when compared to NH children. One adult male and one adult female talker were used. Stelmachowicz et al. found that for the adult male talker's speech, the spectral region 2 – 4 kHz was sufficient to identify these fricatives well, while a wider BW, of 2 – 8 kHz, was required for good identification of the adult female's fricatives. HI children performed on average poorer in their perception of the adult female talker's /s/ and /z/ sounds than they did for the male talker's sounds. HI children showed greater variability in correctly identifying /s/ and /z/ when spoken by a female talker compared to a male talker. The increased errors in /s/ and /z/ produced by the female talker imply poor audibility by the HI child, leading to inconsistent exposure to these important speech sounds.

Stelmachowicz, Lewis, Choi and Hoover (2007) examined the effects of stimulus bandwidth, either 5 kHz or 10 kHz, on fricative perception in nonsense syllables, words and novel words for 7-14 year old NH children and HI children with mild-to-moderate hearing loss. They found, on all tasks, children with HI had poorer performance than children with NH. The results for both groups showed significant BW effects for nonsense syllable and word recognition.

Levine (2010) investigated the effects of talker age and gender on speech perception for normal hearing children and pediatric cochlear implant users. For both open-set word and closed-set vowel (hVd) tests, NH children performed significantly better than pediatric CI users. Also, a significant effect was found for the type of talker, for the pediatric CI users. Specifically, child CI users had poorer scores for words and vowels spoken by young girls than for those spoken by adult males and females.

There appears to be a mismatch between the type of talker frequently used in clinical speech tests (adult males) and the type(s) of talker(s) that children hear most often (adult females and children) (Gilkerson & Richards, 2009). Yet, the perceptual consequences of this mismatch are not entirely known. As described previously, there are differences in the acoustic characteristics of different types of talkers (e.g., more intense energy at high frequencies in children's speech and in females' speech compared to adult males' speech). And, it seems likely that these acoustic differences will yield differences in perceptual results for different types of talkers, especially for listeners with hearing loss using HAs. Important linguistic information in the /s/ sounds produced by young and adult female talkers (6-8 kHz) may not be audible due to limitations in the high frequency gain and output of current HA systems. Results from the many studies by Stelmachowicz and her colleagues, which examine the perception of fricative sounds, hint at a possible effect of talker-type on pediatric listeners' overall speech perception scores. However, this has not been examined explicitly.

Thus, the goal of this study is to determine the effect of talker type (age and gender) on speech perception for children with bilateral mild-moderate sensorineural hearing loss, fit with one or two digital behind-the-ear-hearing aids. Predictions are: 1) across all talker types, HI children will have lower speech perception scores than their NH peers, 2) for HI children, adult male talkers' speech will yield the highest scores, in comparison to those for adult females' and child talkers' speech, 3) speech tests with consonants (which have acoustic energy in much higher spectral regions than vowels) will be more sensitive to talker type effects (i.e., age & gender) than speech tests with vowel sounds.

Methods and Materials

Recruitment letters, informed consent and research protocol for this study were reviewed and approved by the Institutional Review Board (IRB) and the Human Studies Committee at Washington University School of Medicine. Recruitment of participants was completed through collaborators' own patient populations and the Special School District of St. Louis County, with HRPO-approved letters asking for eligible patient referrals. All participants were given a token amount of remuneration for their participation.

Design: This is a prospective, behavioral study.

Subjects: All HA subjects met the following inclusion criteria: chronological age of 7 – 12 years of age, mild-to-moderately-severe bilateral sensorineural hearing loss such that audiometric thresholds in the better ear are less than or equal to 70 dB HL for all test frequencies from 250 to 6000 Hz, currently fit with one and/or two digital behind-the-ear (BTE) hearing aids that had been worn at least 3 months, hearing loss identified by 48 months of age and amplification provided by 5 years of age, native speakers of English, use oral education, and hearing loss as primary disability with normal cognitive function.

Nine hearing-impaired children fit bilaterally with BTE hearing aids met the inclusion criteria and participated in the study. All 9 participants had clear ear canals and tympanic membranes visualized by otoscopy, and tympanometric results that revealed peak pressure and compliance values were within normal limits. Participants (5 females, 4 males) (HA 1-9) ranged in age from 7;1 – 10;8 years;months (mean, 8;8; SD =1;4) and duration of HA use ranged from 3;8 -8;1years;months (mean, 5;8 ; SD = 1.3) (See Table I). All subjects wore bilateral hearing aids; 5 subjects were fit with Phonak, 3 with Widex and 1with Oticon hearing aids; see Table III for detailed device information. No participants were fit with a frequency lowering algorithm (i.e. frequency compression or transposition). Testing was done using the child's own HAs as programmed by their school or clinical audiologist, and no changes were made to the HA prior to or following testing. The frequency-specific gain and output were verified on the Audio Scan Verifit to ensure proper function of each HA. Real-ear-to-coupler differences (RECD) with the child's personal earmolds were used for verifying hearing aid function.

Stimuli: FM (frequency modulated) tones generated by GSI-61 audiometer were used to assess unaided and aided thresholds. Two recorded speech tests spoken by four talker types (adult male, adult female, 10-12 year old girls and 5-7 year old girls) were used to assess speech perception.

Test Protocol: Prior to all speech perception testing, participants underwent otoscopy, tympanomtery, aided threshold testing, and hearing aid verification. Some participants completed unaided threshold testing if a recent audiogram had not been completed and provided by the participant's audiologist within the last 6 months. Aided testing was performed in a single-walled booth with subjects seated approximately 1 meter from an audio speaker at 0 degrees azimuth. All speech material was presented by a Compaq laptop computer routed through a GSI 61 audiometer at a level of approximately 60 dBA at the subject's head. Calibration was performed using a speech-shaped noise signal and a Quest Model 2700 sound level meter.

A modified Hughson-Westlake procedure (Carhart & Jerger, 1959), in increments of 5 dB, was used to obtain both unaided (under supra-aural headphones) and aided thresholds from 250 - 6000 Hz. Each hearing aid was verified prior to speech perception testing using the Audio Scan Verifit system. Real-ear-to-coupler differences (RECDs) were measured unless the participant's audiologist provided RECDs that had been taken in the last 6 months with the patient's current earmolds. Speech maps were obtained at input levels of 50, 60, and 75 dB SPL along with the maximum power output (MPO). The measured output for each level was compared to Desired Sensation Level (DSL 5.0v) targets (Bagatto et al., 2005; Scollie et al., 2005) based on the participants age, transducer used during threshold testing, and binaural fitting of hearing aids. Of the 9 HA users, 7 hit DSL targets; 1 did not meet targets above 2000 Hz for both the left and right HAs and another did not meet targets above 2000 Hz for only the left HA. Aided sound-field thresholds were used to assess minimal audibility from 250-6000 Hz.

The first test, a closed-set vowel test, consisted of eleven vowel sounds spoken in an /h/-vowel-/d/ or (hVd) syllable. The following were target words: "heed, hayed, hid, had, hod, head, heard, hoed, hood, hud, who'd" that correlate appropriately to the vowels /i/, /e/, /I/, /æ/, /a/, /ɛ/, /ɝ/, /o/, /ʊ/, /ʌ/, /u/. Each of the 11 vowel sounds was represented by three tokens.

Across the 11 vowel sounds (33 vowel waveforms), there were twenty-one adult male talkers, twenty-five adult females, fourteen 10-12 year old girls, and three 5-7-year-old girls. The adult males', adult females', and 10-12 year old girls' hVds are a subset of the speech materials recorded by Hillenbrand *et al*. (1995) while the hVds spoken by the 5-7 year old girls are a subset of speech materials originally recorded by Assmann & Katz (2000). Each participant heard 33 hVd syllables (3 tokens of each vowel X 11 vowels) for each of the 4 talker types, for a grand total of 132 hVd different syllables.

For the closed-set vowel test, the CONDOR-ID (Version 2.c) auditory identification program developed by the department of Auditory Implants and Perception at House Ear Institute Los Angeles, CA, was used to present the hVd stimuli and to record responses. Each participant was given a practice run to familiarize him/her with each vowel sound and its corresponding icon on the screen. To ensure that each child was comfortable with the test procedure and familiar with the vowel stimuli, each child was asked to repeat aloud each of the words on the screen when pointed to by the tester. If the child was not comfortable or had incorrect responses, another practice run was completed until the child reliably knew each word and was sure of the task. For the hVd vowel tests, tokens from each talker-type were tested in separate blocks, and within those blocks the 33 vowels were presented in a random order chosen by the CONDOR program. Participants were instructed to click on the corresponding vowel icon on the laptop screen of the word he/she heard.

The second speech perception test was the CNC open-set word lists, developed by Peterson & Lehiste (1962). All 10 CNC word lists contain 50 words, for a total of 150 phonemes. These lists were designed to have speech-sound distributions within each list that approximate the frequency of occurrence of speech sounds found in a large corpus of spoken English (Peterson & Lehiste, 1962). The CNC lists were recorded by four different talker types; adult male, adult female, 11 year-old girl and 6 year-old girl, with one person representing each talker type. Based on the work of Skinner and colleagues (2006), list 9 was omitted from the test battery due to difficulty and intelligibility of the words. CNC list 2 was omitted from the adult female talker and CNC list 10 was omitted from the 10-year-old girl talker (Holden et al., pilot study). Windows media player was used to present the CNC stimuli. Four of the ten available CNC word lists were used; to avoid learning, no list was used more than once. Each participant listened to one list (50 words) of the CNC words per talker type, for a total of 200 CNC words and 600 phonemes.

For the CNC test, the child was instructed to repeat the word presented via the loudspeaker. Each word was preceded by the carrier phrase “ready.” Participants were directed to repeat aloud each word heard after its presentation. The correct word and correct phoneme responses for the CNC lists were scored during testing by the examiner while seated next to the participant in the sound-booth.

The order of the speech tests, by talker type, was randomized and administered as follows: two talker types of CNCs, two talker types of hVds, two talker types of CNCs, and finally two talker types of hVds. Table IV contains the order of presentation of hVd by talker type, Table V contains the order of presentation of CNC word list by talker type, and Table VI provides the CNC list number used for each talker type. A break was given at the halfway point and/or other times throughout the test session if necessary to avoid fatigue or boredom.

Results

Unaided thresholds at 250- 6000 Hz are plotted for each HA subject in Figures 1 and 2 for the left and right ears respectively. In general, most subjects have a greater degree of hearing loss above 1000 Hz (moderate to severe), with a mild to moderate loss at 250-1000 Hz. Across both ears, the range of unaided pure tone thresholds for each frequency are as follows: 250 Hz:

25-55 dB, 500 Hz: 10-65 dB, 1000 Hz: 30-70 dB, 2000 Hz: 45-75 dB, 4000 Hz: 25-80 dB, and 6000 Hz: 45-80 dB.

Bilateral aided thresholds at 250-6000 Hz are plotted for each HA subject in Figure 3. Note the aided thresholds range from 10 to 30 dB HL at 250-2000 Hz and 15 to 50 dB HL at 4000 and 6000 Hz. An examination of hearing aid verification graphs revealed that 7 participants met the gain and output values of DSL targets while 1 participant failed to meet targets above 2000 Hz and above for both HAs and the other participant failed to meet targets above 2000 Hz for only the left HA.

Vowel Identification: For the hVd syllables, group average percent-correct vowel identification scores for each talker-type are shown in Figure 4. Average scores were as follows: adult males (mean = 79%, SD = ±9.6), adult females (mean = 78%, SD = ±12.7), 10-12 year-old girls (mean = 71%, SD = ±13.0) and 5-7 year-old girls (mean = 74%, SD = ±10.8). An ANOVA was performed using the non-parametric Friedman's test since the normality assumption was not met. There was no statistically significant effect of talker-type ($\chi 2 = 5.552$, df = 3, p = 0.136) on the perception of hVd syllables. Table VII shows percent-correct vowel identification scores for each HA listener and each talker type. Vowel confusion matrices are provided in Tables VIIa-VIIId, respectively, for the four talker-types. Table IX shows the overall identification scores for each vowel sound, for each talker type. The vowels identified with the highest accuracy were those in /hid/, /had/ and /heard/, and those identified least accurately were in /who'd/, /hoed/, and /hud/.

CNC words: Group average percent-correct word scores for each talker-type are shown in Figure 5. These HA subjects achieved the highest scores when listening to words spoken by the adult male talker; adult male (mean = 79, SD = ±14), adult female (mean = 66.7%, SD = ±15.7), 10-year old girl (mean = 65.8%, SD = ±7.9) and 6-year old girl (mean = 66.0%, SD = ±11.5). As done for the vowel results, a non-parametric Friedman's ANOVA was performed. Talker-type had a statistically significant effect on CNC word scores ($\chi2 = 12.802$, df = 3, p = 0.005). Post-hoc tests revealed only one significant difference; the male talker's CNCs yielded significantly higher word scores than did the 10-year old girl's CNCs. Table X provides individual data for the CNC words tests for each talker-type.

CNC phonemes: Group average percent-correct phoneme scores for each talker-type are shown in Figure 6. Overall the adult male talker yielded the highest score; adult male (mean = 92.5%, SD = ±6.1), adult female (mean = 85.2%, SD = ±5.3), 10-year old girl (mean = 82.1, SD = ±8.9), and the 6-year old girl (mean = 84.3, SD = ±6.5). Talker-type had a statistically significant effect on CNC phoneme scores ($\chi_2 = 15.635$, df = 3, p = 0.001). Post-hoc tests revealed the male talker's CNCs yielded higher phoneme scores than either the 6-year-old girl's or 10-year-old girl's CNCs. Table X provides individual data on CNC phonemes scores for each talker-type. Table XI displays the most commonly misidentified phonemes for each talker-type, when results are combined across listeners. The most commonly misperceived phoneme was /s/. However, the number of errors in identifying /s/ is not similar for the four talker types; the male talker's CNCs had many fewer errors than found for the other talkers' CNCs. Of the 82 errors for the phoneme /s/, only 5 were made when listening to the adult male; the remaining 77 errors were roughly equally distributed across the three female talkers (adult, 10 and 5 year old). The other commonly misidentified phonemes were /p/, /t/, /f/, /l/, /s/, /d/, /v/, /n/, /b/, and /m/.

Discussion

The effect of talker-type (age and gender) on the speech perception abilities of pediatric HA users is dependent on the type of speech materials used. Specifically, for a closed-set vowel identification task, the talker-type (adult males, adult females, 10-12 year old girls, or 5-7 year old young girls) of the hVd syllables by did not significantly affect percent-correct scores. By contrast, for a CNC word recognition task, talker-type did affect these pediatric bilateral HA users' perceptual scores. Results showed that the average percent-correct word score for the adult male was 79% compared to an average score of 66% for all three females (adult, 10 years old and 6 years old). A similar pattern was obtained for the phoneme scores. Of particular interest, is the disproportionately large number of /s/-errors for the male speaker compared to those for the three female speakers. Nearly all (77 of the 82) the errors in perceiving /s/ were made when these pediatric HA users listened to any of the three female talkers' CNCs.

Acoustically, some of the temporal and spectral differences between adult and children's speech may be attributed to the fact that young children are still developing their speech-motor skills. Thus it seems reasonable that the differences in the perception of adult vs. child talkers may be due, in part, to inaccurate productions by these young talkers. However, differences in perception between male and female adult talkers' speech may be primarily related to the differences in spectral energy for certain sounds, such as /s/ and /sh/, which can have very high-frequency content (>2.5 kHz or so). Indeed, the CNC results for these pediatric HA users are similar to results from Stelmachowicz et al. (2001). They reported that, for NH and HI listeners, there was greater accuracy in perception of the phoneme /s/ when spoken by a male talker than spoken by a female talker or 6-year-old child. An analysis of phoneme errors in this study reveals that a greater proportion of /s/ errors occurred for all three female talker-types than for the male talker. Taken together, these results are consistent in suggesting that the difference in perception of male vs. any female talkers (i.e., adult female or child) may be attributed in part to the limited bandwidth of most hearing aids (i.e., HA bandwidth is often restricted to 4-6 kHz because of attenuated real-ear gain and acoustic feedback limitations for pediatric HA users).

This pattern of results is somewhat different from those found by Levine (2010) for pediatric CI users using these same CNC talkers and speech materials. The child talkers were perceived more poorly than the adult talkers (male and female), and the adult male and adult female talkers were perceived equally well. This may be partially due to the fact that CI devices are able to maintain high-frequency audibility beyond 6 kHz. It is noteworthy that aided thresholds at 4-6 kHz for some of the HA users in this study as large as 50 dB HL, and some children's hearing aids did not meet DSL targets in this frequency range. Additional studies are needed to quantify the possible effects of device bandwidth and talker-type, e.g., examine speech perception abilities of NH children in noise or using low-pass filtered speech to simulate the effects of CIs and HAs, respectively. It may also be worthwhile to explore whether frequency compression or transposition hearing aids could improve perception of female and child talkers by providing an effectively-greater bandwidth than that of typical HAs.

A few factors should be considered, however, before generalizing these results too broadly. A different pattern of results was found in this study for the two different kinds of speech materials, hVds and CNCs. For these pediatric HA users, talker type did not affect the perception of vowels in hVd syllables, while talker type did affect the perception of words and phonemes in CNCs. First, this study employed a relatively small number of participants. Generalizing these results to other pediatric HA users, or to other HA users in general should be done very cautiously. Second, from this study, it is not possible to know whether the difference in pattern of results is due to the types of materials per se (vowels vs. phonetically-balanced word lists) with their associated differences in acoustic (especially spectral) content, or is due to the number of talkers representing each talker-type (multiple talkers for hVds and one talker for CNCs). Recall, the CNC lists were spoken by a single talker for each talker-type, one adult male, one adult female, one 10-year old girl, and one 6-year old girl. Specific acoustic characteristics of an individual talker can affect intelligibility, such that results for these particular talkers may not necessarily generalize to all talkers of the same gender or of similar ages. Before generalizing these CNC results, a larger group of talkers for each talker-type is needed. In addition, incorporating young boy talkers with ages similar to those of the girls used in this study (5-7; 10-12) would beneficial. This would allow an assessment of how children perceive both gender peers, regardless of the similarities of their fundamental frequencies.

Despite the possible limited generality of these results, these results are consistent with those of other studies that found speech perception of HI listeners is affected by talker-type (e.g., Loizou et al., 1998; Fourakis et al., 2004; Fourakis et al., 2007), especially speech perception of young HI listeners (Stelmachowicz et al., 2001; Levine, 2010). Though all these studies employed different listener groups (adults vs. children, hearing impaired vs. normal hearing, CI vs. HA) and different speech stimuli (vowels, CNCs, fricatives), implications from their results for speech testing of young children are clear. Using speech perception tests with adult male speakers may overestimate how well a child performs in everyday communication situations. An over-estimation would occur if children perceive female and child talkers more poorly than male talkers, and if children hear adult females and children more often than they hear adult males. Given the preponderance of females in the roles of caretakers and early educators, it is usually assumed that children hear adult females more often than adult males. However, data to support this notion has been addressed only recently through analyses of natural-setting recordings of spoken language. Gilkerson & Richards (2009) reported that on average, 75% of the ~12800 words spoken daily to children 2 to 30 months old were spoken by females. That is, about 9,600 words were spoken by adult females and 3200 by adult males, to these young children. While, Gilkerson & Richards do not report the number of words spoken by child talkers, they do report that per day, the average 24-month old child produces about 2000 vocalizations “in conversation” with those roughly 12800 words heard from adults. Together, then, adult females’ speech and a child’s own speech constitute a substantial amount of the speech heard by children each day. While it is not yet known whether devices can be optimized for different talker-types, clinicians evaluating and fitting devices for young pediatric cases should consider speech perception tests that are more representative of everyday listening situations.

References

Assmann PF, Katz WF (2000) Time-varying spectral change in the vowels of children and adults. *J. Acoust Soc. Am*, 108: 1856-1866.

Bagatto M, Moodie S, Scollie S, Seewald R, Moodie S, Pumford J, & Liu KPR (2005) Clinical protocols for hearing instrument fitting in the desired sensation level method. *Trends in Amplif*, 9: 199-226.

Boothroyd A (2004) Measuring auditory speech perception capacity in very young children. In R.T. Miyamoto (Ed.), *Cochlear Implants.* International Congress Series 1273 (pp. 292-295). Amsterdam: Elsevier.

Byrne D, Dillon H, Tran K, Arlinger S, Wilbraham K, Cox R...Ludvigsen C (1994) An international comparison of long-term average speech spectra. *J Acoust Soc Am*, 96: 2108- 2120.

Carhart R, Jerger J (1959) Preferred method of clinical determination of pure tone thresholds. *J Speech Hearing Disord*, 24: 330-345.

Eguchi S, Hirsh IJ (1969) Development of speech sounds in children. *Acta Otolaryngol Suppl*, 257: 1-51.

Eisenberg LS, Johnson KC, & Martinez AS (2005) Clinical Assessment of Speech Perception for Infants and Toddlers, *Audiology Online*: ASHA

Gilkerson, J, & Richards, JA (2009) The Power of Talk: Impact of Adult Talk, Conversational Turns, and TV During the Critical 0-4 Years of Child Development (No. ITR-01-2): LENA Foundation.

Hillenbrand J, Getty LA, Clark MJ, & Wheeler K (1995) Acoustic characteristics of American English vowels. *J Acoust Soc Am*, 97: 3099-3111.

Jongman A, Wayland R, & Wong S (2000) Acoustic characteristics of English fricatives. *J Acoustic Soc Am*, 108: 1252-1263.

Kirk KI, Pisoni DB & Miyamoto, RC (1997) Effects of stimulus variability on speech perception in listeners with hearing impairment. *Journal of Speech, Language, and Hearing Research*, 40: 1395-1405.

Kortekaas RW, Stelmachowicz PG (2000) Bandwidth effects on children's perception of the inflectional morpheme /s/: Acoustical measurements, auditory detection, and clarity rating. *Journal of Speech, Language, and Hearing Research*, 43: 645-660.

Lee S, Potamianos A, & Narayanan S (1999) Acoustics of children's speech: developmental changes of temporal and spectral parameters. *J Acoust Soc Am*, 105: 1455-68.

Levine, J (2010) The effect of talker age and gender on speech perception of pediatric cochlear implant users. WUSM Program in Audiology and Communication Sciences (Doctorate of Audiology Capstone Report), St. Louis, MO.

Peterson GE, Lehiste I (1962) Revised CNC lists for auditory tests. *J Speech Hearing Disord,* 27: 62-70.

Pittman AL, Stelmachowicz PG, Lewis DE, & Hoover BM (2003) Spectral characteristics of speech at the ear: Implications for amplification in children. *Journal of Speech, Language, and Hearing Research*, 46; 649-657.

Scollie S, Seewald R, Cornelisse L, Moodie S, Bagatto M, Laurnagaray D, Beaulac S, & Pumford J (2005) The desired sensation level multistage input/output algorithm. *Trends in Amplif,* 9: 159-197.

Skinner MW, Miller JD (1983) Amplification band-width and intelligibility of speech in quiet and noise for listeners with sensorineural hearing loss. *Audiology,* 22: 253-279.

Sommers MS, Kirk KI & Pisoni DB (1997) Some considerations in evaluating spoken word recognition by normal-hearing, noise-masked normal-hearing, and cochlear implant listeners. I: The effects of response format. *Ear Hear,* 18: 89-99.

Stelmachowicz PG, Pittman AL, Hoover BM, & Lewis DE (2001) Effect of stimulus bandwidth on the perception of /s/ in normal- and hearing-impaired children and adults. *J Acoust Soc Am*, 110: 2183-90.

Stelmachowicz PG, Pittman AL, Hoover BM, & Lewis DE (2002) Aided perception of /s/ and /z/ by hearing impaired children. *Ear Hear*, 23: 316–324.

Stelmachowicz PG, Lewis DE, Choi S, & Hoover BM (2007) Effect of stimulus bandwidth on auditory skills in normal-hearing and hearing-impaired children. *Ear Hear,* 28: 483-94.

Tables

Table I: Demographic information on HA subjects. Age = years; months

Subject	Age	Gender	Age at hearing loss identification	Age at amplification	Etiology
HA1	7;4	F	3;6	3;6	EVA
HA2	8;6	M	0	0;5	Unknown
HA3	9;4	F	2;0	3;0	Unknown
HA4	7;10	F	4;0	4;0	Unknown
HA5	9;2	M	4;0	4;0	Unknown
HA6	10;7	M	4;6	4;6	Unknown
HA7	7;6	M	0	0;8	Unknown
HA8	7;1	F	1;3	2;6	EVA
HA9	10;8	F	4;0	5;0	Unknown

Table II: Hearing Aid Make and Model used by HA subjects.

Subject	Left HA	Right HA
HA1	Phonak Naida III UP	Phonak Naida III UP
HA2	Phonak Savia	Phonak Savia
HA3	Widex Inteo-19	Widex Inteo - 19
HA4	Oticon Tego Pro	Oticon Tego Pro
HA5	Widex Inteo – 9	Widex Inteo - 19
HA6	Widex Senso Diva – 9	Widex Senso Diva - 9
HA7	Phonak Naida V SP	Phonak Naida V SP
HA8	Phonak Eleva	Phonak Eleva
HA9	Phonak Naida III SP	Phonak Naida III SP

Table III: Presentation order for hVd vowels by talker type.
(1 = Adult Males, 2 = Adult Females, 3 = 10-12 year old girls, 4 = 5-7 year old girls)

Subject	1st	2nd	3rd	4th
HA1	3	1	2	4
HA2	2	1	4	3
HA3	4	1	3	2
HA4	1	2	3	4
HA5	4	2	3	1
HA6	2	4	1	3
HA7	3	1	4	2
HA8	1	3	4	2
HA9	4	3	2	1

Table IV: Presentation order for CNC words by talker type for CNC.
(1= Adult Male, 2= Adult Female, 3= 10-year old girl, 4 = 6-year old girl)

Subject	1st	2nd	3rd	4th
HA1	1	3	2	4
HA2	2	3	1	4
HA3	2	1	3	4
HA4	3	1	4	2
HA5	4	2	3	1
HA6	1	4	3	2
HA7	1	2	4	3
HA8	4	1	2	3
HA9	2	3	4	1

Table V: CNC list number administered to each HA subject per talker type.
Lists omitted are: List #9 for Adult Male, Lists #2 & #9 for Adult Female, List #9 & #10 for the 10-year old girl, and List #9 for the 6-year old girl.

Subject	Adult Male	Adult Female	10-year old girl	6-year old girl
HA1	4	8	1	3
HA2	6	4	3	5
HA3	7	3	2	1
HA4	6	5	7	2
HA5	3	8	6	4
HA6	5	1	7	10
HA7	2	3	1	7
HA8	10	1	3	6
HA9	7	6	8	4

Table VI: Vowel identification scores (% correct) for hVds (N = 33 trials per cell: 11 vowels x 3 presentations per vowel for each talker type).

Subject	Adult Males	Adult Females	10-12 year old girls	5-7 year old girls	Avg
HA 1	82	85	70	79	78.8
HA 2	82	88	85	88	85.6
HA 3	91	76	88	58	78.0
HA 4	61	67	58	73	64.4
HA 5	67	70	48	67	62.9
HA 6	88	76	73	91	81.8
HA 7	82	94	82	76	83.3
HA 8	76	55	73	64	66.7
HA 9	82	91	64	70	76.5
Average	78.8	77.8	71.0	73.7	

Tables VIIa-VIId: Vowel confusion matrices for the hVd syllables, combined responses across subjects HA1-HA9. One confusion matrix per talker type. Entries are counts.

VIIa: Adult Male talkers

Subject	Heed	Hid	Head	Had	Hod	Hood	Who'd	Hud	Heard	Hayed	Hoed
Heed	**20**	1	2								
Hid	6	**25**	2							3	
Head	1		**23**			1				4	
Had				**27**	2						
Hod					**22**	1	3	5			10
Hood		1			1	**22**	4	7	1		1
Who'd							**19**				1
Hud					1	3		**15**			
Heard									**26**		
Hayed										**20**	
Hoed					1		1				**15**
Presentations	27	27	27	27	27	27	27	27	27	27	27
Responses	23	36	29	29	41	37	20	19	26	20	20

VIIb: Adult Female talkers

Subject	Heed	Hid	Head	Had	Hod	Hood	Who'd	Hud	Heard	Hayed	Hoed
Heed	**22**	1	2								
Hid	5	**25**									
Head		1	**24**		2					2	
Had				**27**	6			1		1	
Hod					**17**		4	3			7
Hood						**23**	6	1	2		3
Who'd							**14**				1
Hud			1		1	3	1	**22**	1		1
Heard						1			**24**		1
Hayed										**24**	
Hoed					1		2				**14**
Presentations	27	27	27	27	27	27	27	27	27	27	27
Responses	21	32	28	38	27	40	9	31	25	25	21

VIIc: 10-12 year old girl talkers

Subject	Heed	Hid	Head	Had	Hod	Hood	Who'd	Hud	Heard	Hayed	Hoed
Heed	**17**										
Hid	7	**26**	1			1				12	
Head	2		**21**	2				1			
Had			5	**22**	3			1		1	
Hod	1			2	**20**		2	2			10
Hood		1			1	**21**	11		2	1	3
Who'd							**11**				
Hud				1	2	5	1	**22**			
Heard					1				**25**		1
Hayed										**13**	
Hoed							2	1			**13**
Presentations	27	27	27	27	27	27	27	27	27	27	27
Responses	17	47	26	32	37	37	11	31	27	13	16

VIId: 5-7 year old girl talkers

Subject	Heed	Hid	Head	Had	Hod	Hood	Who'd	Hud	Heard	Hayed	Hoed
Heed	**22**										
Hid	1	**24**	6			2				1	
Head	2	3	**17**	1		1		2			
Had			3	**25**						2	
Hod					**26**		1	7			5
Hood	1		1			**22**	9	5			8
Who'd							**9**				
Hud					1		3	**13**	3		
Heard						2	1		**24**		1
Hayed	1			1			1			**24**	
Hoed							3				**13**
Presentations	27	27	27	27	27	27	27	27	27	27	27
Responses	22	35	25	30	40	42	9	20	28	27	16

Table VIII: Average percent-correct identification scores, per vowel, combined across subjects HA1-HA9, for each talker type. (27 trials per cell in table; 9 subjects x 3 presentations per talker type).

Talker Type	Heed	Hid	Head	Had	Hod	Hood	Who'd	Hud	Heard	Hayed	Hoed	Average
Adult Males	74	93	85	100	73	73	63	50	87	67	50	74
Adult Females	81	93	89	100	63	85	52	81	89	89	52	79
10-12 year old Girls	63	96	78	81	74	78	41	81	93	48	48	71
5-7 year old Girls	81	89	63	93	96	81	33	48	89	89	48	74
Average	75	93	79	94	77	79	47	65	90	73	50	

Table IX: Identification scores for CNCs. (N = 50 words; N = 150 phonemes)
(% correct words; % correct phonemes)

Subject	Male	Female	10 G	6 G	Avg
HA 1	86;97	66;87	62;85	72;90	72;90
HA 2	80;92	64;79	60;62	52;76	64;77
HA 3	76;91	68;86	64;85	68;84	69;87
HA 4	80;93	64;85	72;87	72;86	72;88
HA 5	80;93	66;86	66;86	52;75	66;85
HA 6	90;96	68;89	82;92	86;95	82;93
HA 7	78;91	90;95	68;85	70;89	77;90
HA 8	64;88	54;77	54;74	68;81	60;80
HA 9	80;92	60;83	64;83	54;83	65;85
Average	79;93	67;85	66;82	66;84	

Table X: The ten most commonly misidentified consonants from the CNC word tests. Number of errors combined across subjects HA1 – HA9.

	Consonant									
Talker-Type	S	P	T	D	L	N	F	B	V	M
Male	5	12	12	6	6	5	9	6	3	3
Female	27	25	5	7	16	15	8	12	9	9
10-year old Girl	25	24	21	19	6	9	11	4	4	7
6-year old Girl	25	11	16	13	11	10	8	8	10	7
Sum	82	72	54	45	39	39	36	30	26	26

Figure 1: Unaided pure-tone thresholds for the left ear under headphones/inserts using conventional audiometry from 250-6000 Hz.

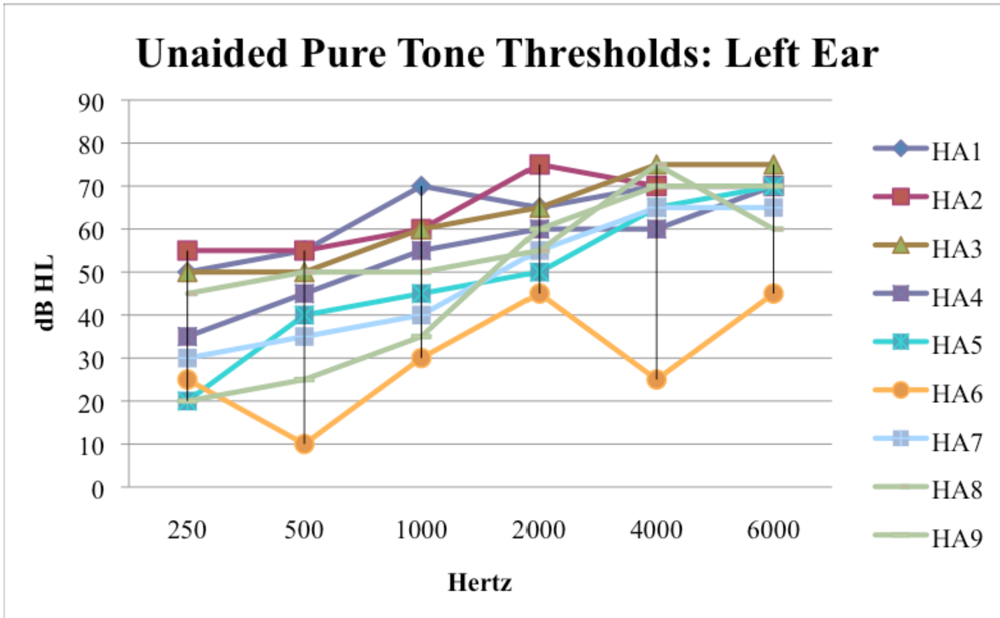

Figure 2: Unaided pure-tone thresholds for the right ear under headphones/inserts using conventional audiometry from 250-6000 Hz.

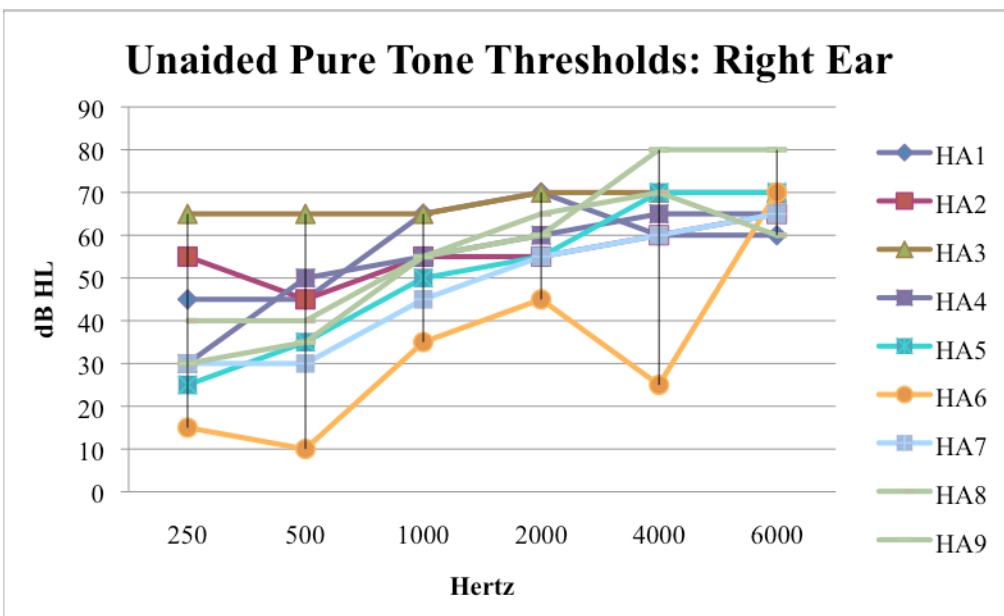

Figure 3: Aided bilateral audiometric thresholds performed in the soundfield using conventional audiometry from 250-6000 Hz.

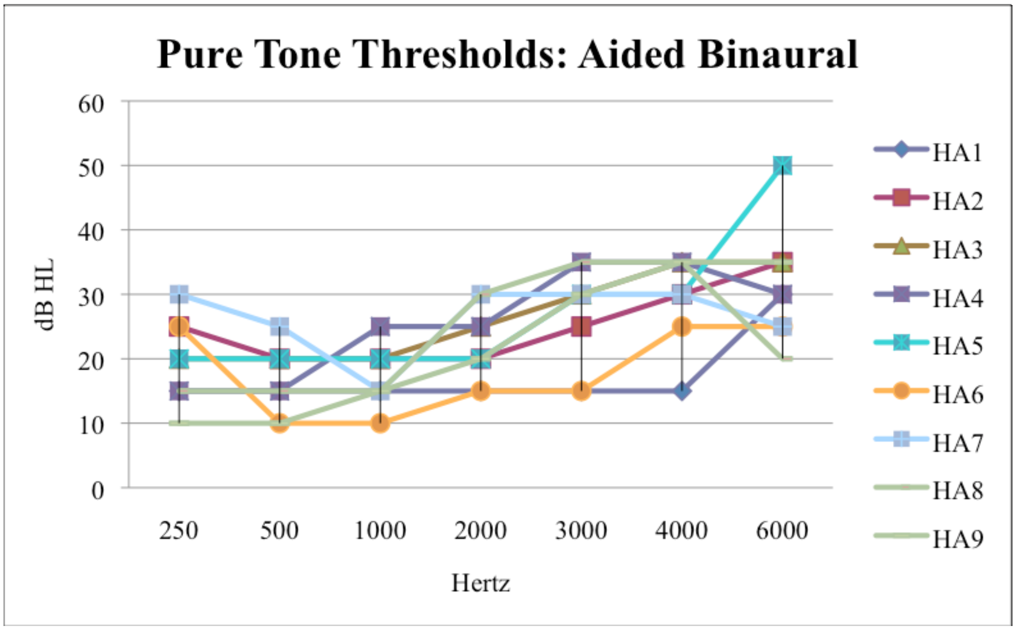

Figure 4: Average hVd scores across HA1-HA9 subjects, +/- 1 SD.

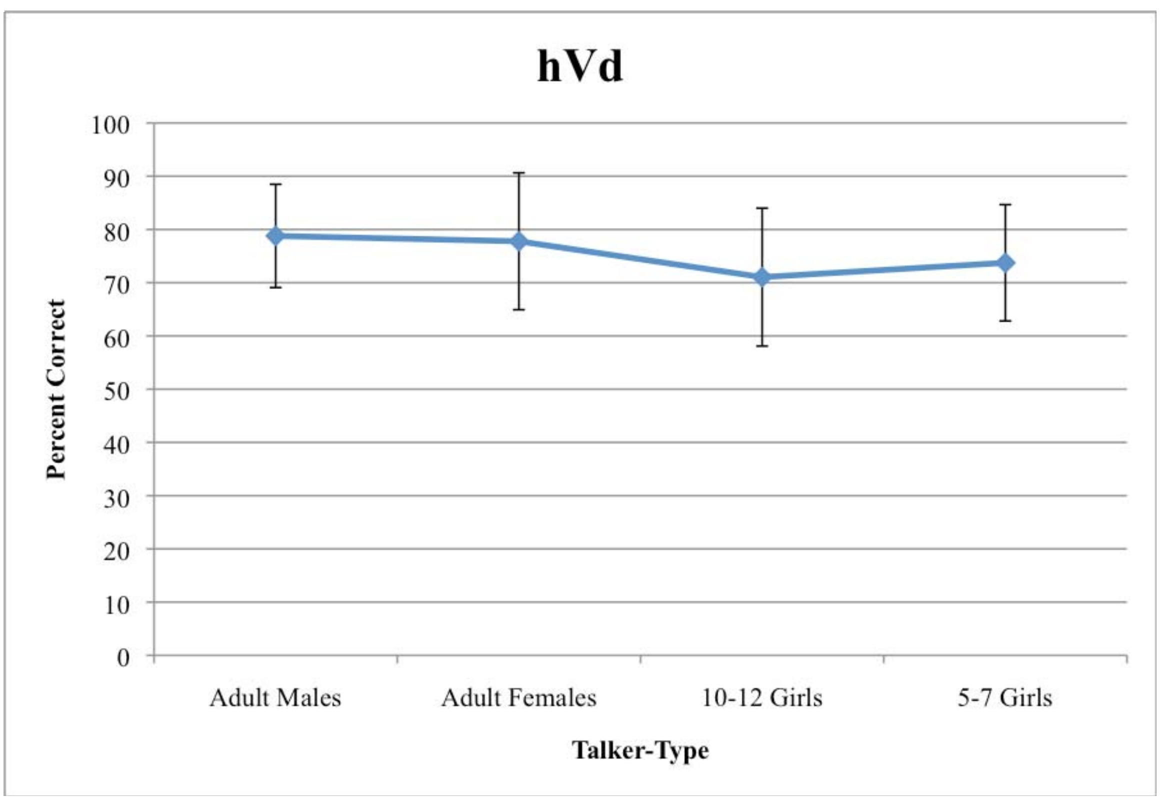

Figure 5: Average CNC scores across HA1-HA9 subjects, +/- 1 SD.

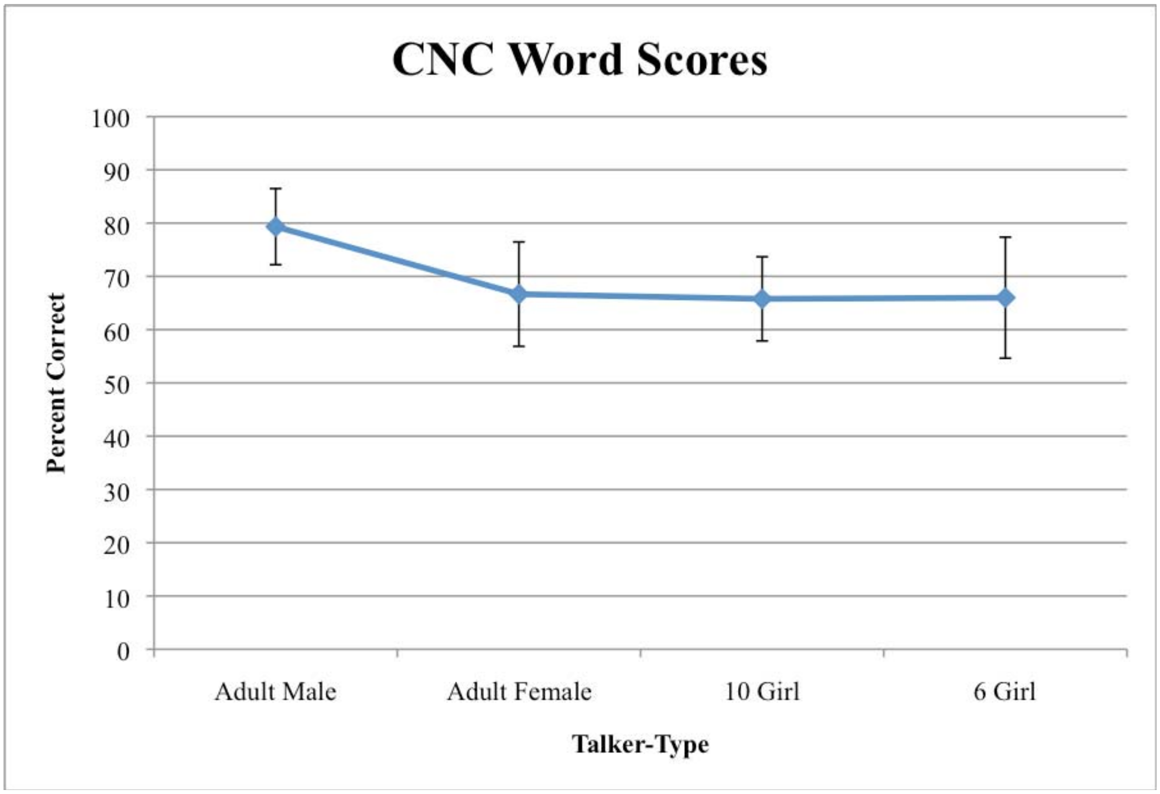

Figure 6: Average CNC phonemes scores from HA1-HA9 subjects, +/- 1 SD.