# Audibility-based predictions of speech recognition for children and adults with normal hearing

Ryan W. McCreery[a] and Patricia G. Stelmachowicz

*Boys Town National Research Hospital, 555 North 3th Street, Omaha, Nebraska 68131*

This study investigated the relationship between audibility and predictions of speech recognition for children and adults with normal hearing. The Speech Intelligibility Index (SII) is used to quantify the audibility of speech signals and can be applied to transfer functions to predict speech recognition scores. Although the SII is used clinically with children, relatively few studies have evaluated SII predictions of children's speech recognition directly. Children have required more audibility than adults to reach maximum levels of speech understanding in previous studies. Furthermore, children may require greater bandwidth than adults for optimal speech understanding, which could influence frequency-importance functions used to calculate the SII. Speech recognition was measured for 116 children and 19 adults with normal hearing. Stimulus bandwidth and background noise level were varied systematically in order to evaluate speech recognition as predicted by the SII and derive frequency-importance functions for children and adults. Results suggested that children required greater audibility to reach the same level of speech understanding as adults. However, differences in performance between adults and children did not vary across frequency bands. © *2011 Acoustical Society of America.* [DOI: 10.1121/1.3658476]

## I. INTRODUCTION

The primary goal of amplification for children with permanent hearing loss is to restore audibility of the speech signal to facilitate development of communication (Bagatto *et al.*, 2010; Seewald *et al.*, 2005). Quantifying audibility is therefore essential to ensure that children have sufficient access to the acoustic cues that facilitate speech and language development. Because speech recognition is difficult to assess with infants and young children, clinicians must rely on indirect estimates of speech understanding derived from acoustic measurements of the hearing-aid output. These predictions are based on adult speech recognition data, despite extensive evidence that suggests that speech recognition in children cannot be accurately predicted from adult data. The purpose of the current study is to evaluate the efficacy of using adult perceptual data to estimate speech understanding in children.

A common method used to quantify the audibility of speech is the Speech Intelligibility Index (SII) (ANSI, 1997). The SII specifies the weighted audibility of speech across frequency bands. Calculation of the SII requires the hearing thresholds of the listener and the spectrum level of both the background noise and speech signals. For each frequency band, a frequency-importance weight is applied to estimate the contribution of that band to the overall speech recognition score. Additionally, the audibility of the signal is determined by the level of the speech spectrum compared to either the listener's threshold or noise spectrum in each band, whichever is greater. The SII is calculated as

$$SII = \sum_{i=1}^{n} I_i A_i,$$ (1)

where $n$ represents the number of frequency bands included in the summation. $I_i$ and $A_i$ represent the importance and audibility coefficients for each frequency band (i), which are multiplied and summed to produce a single value between 0 and 1. An SII of 0 indicates that none of the speech signal is audible to the listener, whereas an SII of 1 represents a speech signal that is fully audible.

In addition to providing a numeric estimate of audibility, SII values can be compared to transfer functions in order to estimate speech recognition for a specific level of audibility. Transfer functions based on adult listeners with normal hearing have been empirically derived for speech stimuli ranging in linguistic complexity from nonword syllables to sentences. Transfer functions have steeper slopes and asymptote at lower SII values when stimuli contain lexical, semantic or syntactic cues, reflecting the listener's ability use of linguistic and contextual information when the audibility of acoustic cues is limited (Pavlovic, 1987). Transfer functions can be used to estimate the effect of limited audibility on speech understanding or determine the amount of audibility required to achieve optimal performance. The impact of limited audibility on speech understanding of young children who may be unable to participate in speech recognition testing due to their age and/or level of speech and language development could be more accurately estimated from SII transfer functions derived from children than using data from adults.

The accuracy of speech recognition predictions based on the SII for adult listeners with normal hearing and hearing loss has been demonstrated and refined in multiple studies over the last 65 years (French and Steinburg, 1947; Fletcher

---

[a]Author to whom correspondence should be addressed. Electronic mail: ryan.mccreery@boystown.org

and Galt, 1950; Dubno *et al.*, 1989; Ching *et al.*, 1998; Ching *et al.*, 2001; Humes, 2002). Despite the potential utility of the SII for estimating audibility and predicting speech recognition in children, the accuracy of SII predictions for children has not been extensively studied. Stelmachowicz *et al.* (2000) measured speech understanding with sentences in noise for children with normal hearing, children with hearing loss, and adults with normal hearing. Results indicated that in order to reach levels of speech recognition similar to adults, both groups of children required greater audibility as measured by the SII. Other studies have evaluated the relationship between audibility and speech understanding in children to determine the stimulus characteristics that may predict the need for greater audibility in children. Broader stimulus bandwidths (Stelmachowicz *et al.*, 2001; Mlot *et al.*, 2010) and higher sensation levels (Kortekaas and Stelmachowicz, 2000) have been found to maximize speech recognition in children. Collectively, these studies suggest that children require more audibility than adults to achieve similar levels of performance and that some of these differences may be dependent on stimulus bandwidth. If children require greater bandwidth for optimal speech understanding, frequency-importance weights for high-frequency bands may be distributed differently than weights obtained with adults. Because importance weights represent the degradation in speech understanding that occurs when a specific band is removed from the stimulus, differences in speech recognition between adults and children that are related to bandwidth are likely to influence the calculation of the SII.

Accurate predictions of speech recognition based on the SII are also dependent on the stability of frequency-importance weights across conditions of varying audibility. Gustafson and Pittman (2011) evaluated this assumption by varying the sensation level and bandwidth of meaningful and nonsense sentences to create conditions with equivalent SII, but different bandwidths. If the contribution of each frequency band is independent of sensation level, equivalent performance should have been observed across conditions with the same SII. However, for conditions with equivalent SII, speech recognition was higher in conditions with a lower sensation level and a broader bandwidth than in conditions where bandwidth was restricted but presented at a higher level. These results demonstrate that the loss of frequency bands cannot be compensated for by increasing sensation level, and that listening conditions with equivalent SII may not necessarily result in similar speech understanding.

Despite the emergence of studies evaluating the effects of audibility on speech recognition with children, few studies have directly assessed variability in SII predictions of speech recognition with children. Scollie (2008) examined this relationship for normal-hearing children and adults, and children with hearing loss. Using nonsense disyllables, speech recognition was measured under conditions with varying audibility. Both groups of children had poorer speech recognition than adults for conditions with the same audibility. Differences in speech understanding of 30% or greater were observed at the same SII, both within groups of children and between children and adults. These findings demonstrate that the SII is likely to overestimate speech understanding for children

and does not reflect the variability of children's speech recognition skills. The limitations of the SII to predict speech recognition in children have significant implications for the clinical utility of these measures. Specifically, the consequences of limited bandwidth or noise on speech understanding are likely to be underestimated for children when the SII is used as the reference.

Because speech recognition data are the basis for the SII, differences between the auditory skills of adults and children are likely to affect both the transfer functions and frequency-importance weights used to quantify audibility and predict speech understanding. Transfer functions that relate audibility to speech recognition scores have been primarily developed using adults. Speech recognition in conditions with background noise, reverberation or limited bandwidth improves as a function of age throughout childhood (Elliot, 1979; Nabalek and Robinson, 1982; Johnson, 2000; Neuman *et al.*, 2010; Stelmachowicz *et al.*, 2001). Performance-intensity functions obtained with children suggest that children's speech recognition is more variable and requires a higher SNR to reach maximum performance than adults (McCreery *et al.*, 2010). While audibility is necessary for speech understanding in children, audibility alone does not appear to be sufficient to predict the range and variability in speech recognition outcomes in children. Particularly for children who vary in their linguistic and cognitive skills, audibility is only one of the potential factors that could influence speech recognition. Children's speech recognition has been shown to depend on immediate memory (Eisenberg *et al.*, 2000) and linguistic knowledge (Hnath-Chisolm *et al.*, 1998). Therefore, age-specific transfer functions may be more accurate than adult-derived estimates.

Frequency-importance weights in the SII calculation could also differ between adults and children. Frequency-importance functions numerically quantify the importance of each frequency band to the overall speech recognition score for a specific stimulus. These functions have been derived from adult listeners for a wide range of varying speech stimuli, including continuous discourse (Studebaker *et al.*, 1987), high- and low-context sentences (Bell *et al.*, 1992), monosyllabic word lists such as CID W-22 (Studebaker and Sherbecoe, 1991), NU-6 (Studebaker *et al.*, 1993), and nonsense words (Duggirala *et al.*, 1988). Overall, these studies reveal that for stimuli with redundant linguistic content such as familiar words or sentences, importance weights for adult listeners are shifted towards the mid-frequencies and reduced at higher frequencies. For speech stimuli with limited linguistic content, such as nonsense words, adult listeners exhibit frequency importance that is spread more evenly across frequency bands with larger importance weights for high-frequency bands. This pattern suggests that when linguistic and contextual information is limited, listeners require more spectral information, particularly at frequencies above 3 kHz, for accurate speech recognition. Because children are developing language, frequency-importance weights for children may be distributed more evenly across bands and larger for high-frequency bands, similar to adults for stimuli with limited linguistic context. Children experience greater degradation in speech understanding when high-frequency

spectral cues are limited (Eisenberg *et al.*, 2000), and differences in how children use high-frequency spectral cues to facilitate speech recognition may alter the distribution of frequency-importance weights for young listeners. Differences in importance weights could lead to different estimates of audibility and speech recognition for children, particularly for amplification where the audibility of the frequencies above 4 kHz may be limited by hearing loss, hearing-aid bandwidth or other factors (Moore *et al.*, 2008). However, because importance weights have not been derived previously for children, SII-based audibility and related predictions of speech recognition for children are calculated using adult weights.

Multiple challenges are likely to have prevented previous attempts to obtain frequency-importance functions with children. Frequency-importance weights are derived using a speech recognition task in which the speech signal is high- and low-pass filtered to systematically evaluate the relative contribution of each frequency region to the overall recognition score. The importance weight for each band is determined by the average amount of degradation in speech recognition that occurs when a band is removed from the stimulus for a large group of listeners (Fletcher and Galt, 1950; Studebaker and Sherbecoe, 1991). The number of frequency bands used with adults varies from a minimum of six bands for the octave-band calculation procedure to 21 bands for the critical-band method. Additionally, speech recognition is measured at multiple SNRs for each frequency band to estimate the contribution of a specific band over a range of audibility. As a result, studies of adult frequency-importance functions may have more than 60 listening conditions due to the combinations of bandwidth and SNR conditions that must be assessed. Even if the task were adapted to limit the number of conditions to avoid age-related confounds such as limited attention and fatigue, the minimum number of listening conditions that would be required if four SNRs are used for an octave band method would be approximately 28 once a full bandwidth condition is included.

The linguistic complexity of the stimuli used to obtain frequency-importance weights with children is also an important experimental consideration that could significantly influence the importance values obtained from the task. Because importance weights derived from adult listeners show varying levels of importance based on the availability of lexical, semantic, syntactic and other linguistic cues in the stimuli, the listener's knowledge and ability to use these cues is likely to influence frequency-importance functions for different speech stimuli. Although not specifically examined in previous studies of importance functions with adults, the probability of occurrence of combinations of phonemes, or phonotactic probability, of the stimuli has been shown to influence the ability of children to identify nonword speech tokens (Munson *et al.*, 2005; Edwards *et al.*, 2004). While previous studies have demonstrated that children as young as 4 years of age are able to use linguistic cues to support speech recognition under adverse conditions (Nittrouer and Boothroyd, 1990), children are likely to vary in their ability to use these cues. Variability in speech recognition ability for children, even within the same age group, presents

challenges in the development of valid frequency-importance functions for this population.

The goal of the present study was to evaluate age-related changes in the relationship between the SII and speech recognition. To further our understanding of the mechanisms underlying these differences, age- and frequency-dependent differences in speech recognition were evaluated by deriving frequency-importance functions from children and adults with normal hearing. A modified filtered speech recognition paradigm that has been used in previous frequency-importance studies with adults (Studebaker and Sherbecoe, 1991, 2002) was used to test three hypotheses in children between 5 and 12 years of age and a group of young adults. First, speech recognition for younger children was expected to be poorer and more variable than for adults and older children when compared across listening conditions with the same SII. Second, based on previous studies that suggest that children require greater bandwidth to reach the same level of speech recognition as adults, the distribution of weights across frequency bands was expected to vary as a function of age. Specifically, children would be expected to have larger weights than adults for the 4 kHz and 8 kHz bands, reflecting larger degradation in performance when these bands are removed from the stimulus. This difference was expected to decrease for older children. Previous studies have found that children may experience poorer perception of fricative sounds than adults as high-frequency audibility is limited (Stelmachowicz *et al.*, 2001). Other studies suggest that age-related differences in speech recognition are related to general perceptual differences that result in increased consonant error rates across multiple categories of phonemes (Nishi *et al.*, 2010). Analyses of consonant error patterns were completed as a function of age and filter condition to characterize potential developmental patterns of variability.

## II. METHOD

### A. Participants

137 individuals participated in the current study. 116 children between 5 years, 3 months and 12 years, 11 months [mean age $= 9.16$ years, standard deviation (SD) $= 2.13$ years] participated, which included thirteen 5 year-olds, thirteen 6 year-olds, fourteen 7 year-olds, sixteen 8 year-olds, fourteen 9 year-olds, nineteen 10 year-olds, fourteen 11 year-olds, and thirteen 12 year-olds. The study also included nineteen adults between 20 and 48 years [mean $= 29.9$ years, SD $= 8.53$ years]. All participants were recruited from the Human Research Subjects Core database at Boys Town National Research Hospital. The participants were paid $12 per hour for their participation, and children also received a book. All listeners had clinically normal hearing in the test ear (15 dB HL or less) as measured by pure-tone audiometry at octave frequencies from 250 Hz to 8000 Hz. One child and one adult did not meet the audiological criteria for the study and were excluded. None of the participants or their parents reported any history of speech, language or learning problems. Children were screened for articulation problems that could influence verbal responses using the Bankson Bernthal Quick Screen of Phonology (BBQSP) (Bankson

and Bernthal, 1990). The BBQSP is a clinical screening test that uses pictures of objects to elicit productions of words containing target phonemes. One child did not pass the age-related screening criterion and was excluded. Expressive language skills were measured for each participant using the Expressive Vocabulary Test, Second Edition, Form A (EVT-2) (Williams, 2007). All of the children in the study were within two SD of the normal range for their age [Mean = 105; Range = 80−121].

## B. Stimuli

Consonant-vowel-consonant (CVC) nonword stimuli were developed for this study. The stimuli were created by taking all possible combinations of CVC using the consonants /b/, /tʃ/, /d/, /f/, /g/, /h/, /dʒ/, /k/, /m/, /n/, /p/, /s/, /ʃ/, /t/, /θ/, /ð/, /v/, /z/, and /ʒ/ and the vowels /a/, /i/, /I/, /ɛ/, /u/, /ʊ/, and /ʌ/. The resulting CVC combinations were entered into an online database based on the Child Mental Lexicon (CML) (Storkel and Hoover, 2010) to identify all of the CVC stimuli that were real words likely to be within a child's lexicon and to calculate the phonotactic probability of each nonword using the biphone sum of the CV and VC segments. All of the real words and all of the nonwords that contained any biphone combination that was illegal in English (biphone sum phonotactic probability = 0) were eliminated. Review of the remaining CVCs was completed to remove slang words and proper nouns that were not identified by the calculator. After removing all real words and phonotactically illegal combinations, 1575 nonword CVCs remained. In order to create a set of stimuli with average phonotactic probabilities, the mean and SD of the biphone sum for the entire set was calculated. In order to limit the variability in speech understanding across age groups, the 735 CVC nonwords with phonotactic probability within ±0.5 SD from the mean were included. Stimuli were recorded for two female talkers at rate of 44.1 kHz.

Three exemplars of each CVC nonword were recorded. Two raters independently selected the best production of each CVC on the basis of clarity and vocal effort. In 37 cases where the two raters did not agree, a third rater listened to the nonwords and blindly selected the best production using the same criteria. To ensure that the stimuli were intelligible, speech recognition testing was completed with three adults with normal hearing. Stimuli were presented monaurally at 60 dB SPL under Sennheiser HD-25-1 headphones. Any stimulus that was not accurately repeated by all three listeners was excluded. Finally, the remaining words (725) were separated into 25-item lists that were balanced for occurrence of initial and final consonant.

Stimulus filtering was completed in MATLAB using a series of infinite-impulse response Butterworth filters, as in previous studies (Studebaker and Sherbecoe, 2002). Removing each octave band through filtering resulted in three high-pass and three low-pass conditions that corresponded with the center frequencies specified by the octave-band method for the SII (ANSI, 1997). The filter slope varied across conditions but was greater than 200 dB/octave for all conditions. Table I displays the filter bandwidths for each condition.

TABLE I. Filter conditions and frequency bands.

| Condition | Frequency range |
| --- | --- |
| Full band (FB) | 88–11 000 Hz |
| Low-pass (LP) | |
| LP1 | 88–5657 Hz |
| LP2 | 88–2829 Hz |
| LP3 | 88–1415 Hz |
| High-pass (HP) | |
| HP1 | 354–11 000 Hz |
| HP2 | 707–11 000 Hz |
| HP3 | 1415–11 000 Hz |

Steady-state speech-shaped masking noise was created to match each talker's long-term average speech spectrum (LTASS). Figure 1 shows the LTASS for each talker compared to the LTASS used in the ANSI standard for SII calculation. The long-term average spectra for the female talkers in the current study were analyzed using concatenated sound files of each talker presented at 60 dB SPL through the Sennheiser HD-25-1 headphones coupled to a Larson Davis (LD) System 824 sound level meter with a LD AEC 101 IEC 318 headphone coupler. The recording of each talker was analyzed electronically using custom MATLAB software using an Fast Fourier Transform (FFT) with a 40 ms window and 20 ms overlap. The spectrum of the ANSI standard is based on a male talker, whereas the two talkers for the current study were both female.

The steady-state masking noise based on the average LTASS for the two female talkers was created in MATLAB by taking an FFT of a concatenated sound file containing all of the stimuli produced for each talker, randomizing the phase of the signal at each sample point, and then taking the inverse FFT. This process preserves the long-term average spectrum, but eliminates temporal and spectral dips.

## C. Instrumentation

Stimulus presentation and response recording was performed using custom software on a personal computer with a Lynx Two-B sound card. Sennheiser HD-25-1 headphones were used for stimulus presentation. A Shure BETA 53 head-worn boom microphone connected to a Shure M267 microphone amplifier/mixed was used to record subject
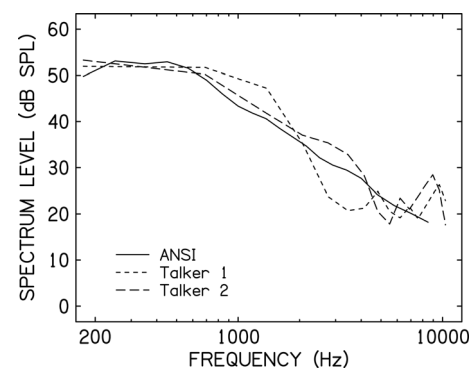


FIG. 1. Long-term average speech spectrum as a function of frequency normalized to 60 dB SPL with ANSI standard (solid line- S3.5-1997), Talker 1 (small dashed line) and Talker 2 (large dashed line).

responses. Pictures were presented via a computer monitor during the listening task to maintain subject interest. The levels of the speech and noise were calibrated using the LD System 824 sound level meter with a LD AEC 101 IEC 318 headphone coupler. Prior to data collection for each subject, levels were verified by measuring a pure tone signal using a voltmeter and comparing the voltage to that obtained during the calibration process for the same pure tone.

## D. SII calculations

For each combination and filter condition, the SII was calculated. The octave-band method was used with the ANSI frequency-importance weighting function for nonsense syllables and a non-reverberant environment. The octave band spectrum levels of the speech and noise stimuli were measured using the same apparatus used for calibration. The levels of speech and noise were converted to free-field using the eardrum to free-field transfer function from the SII. The octave band spectrum levels of speech and noise for each condition were used to calculate the SII for each condition. The SII for each combination of filter and SNR are listed in Table II.

## E. Procedure

Participants and the parents of children who participated took part in a consent/assent process as approved by the local Institutional Review Board. All of the procedures were completed in a sound-treated audiometric test room. Pure tone audiometric thresholds were obtained using TDH-49 earphones. The children completed the BBQSP and EVT. Participants were seated at a table in front of the computer monitor and instructed that they would hear lists of words that were not real words and to repeat exactly what they heard. Participants were encouraged to guess. Each subject completed a practice trial in the full bandwidth condition at the most favorable SNR (+9 dB) to ensure that the task and directions were understood. To obtain a baseline under more favorable conditions, the full bandwidth and LP1 conditions were completed in quiet.

Following completion of the practice trial and two quiet conditions, the filtered speech recognition task was completed in noise using one 25-item list per condition. List number, talker, filter condition and SNR were randomized.

TABLE II. Speech intelligibility index calculations for each combination of filter and SNR.[a]

| | SNR | | | | |
|---|---|---|---|---|---|
| Bandwidth | 0 dB | 3 dB | 6 dB | 9 dB | Quiet |
| FB | 0.41 | 0.49 | 0.57 | 0.66 | 0.99 |
| LP1 | 0.39 | 0.47 | 0.55 | 0.63 | 0.88 |
| LP2 | 0.29 | 0.36 | 0.41 | 0.47 | — |
| LP3 | 0.16 | 0.19 | 0.23 | 0.26 | — |
| HP1 | 0.39 | 0.47 | 0.54 | 0.62 | — |
| HP2 | 0.34 | 0.41 | 0.47 | 0.54 | — |
| HP3 | 0.25 | 0.30 | 0.35 | 0.40 | — |

[a]SII = Speech intelligibility index, SNR = signal-to-noise ratio, FB = full bandwidth, LP = low-pass, HP = high-pass.

The presentation order of the stimuli within each list was also randomized. Although feedback was not provided on a trial-by-trial basis, children were encouraged regardless of their performance after each list. Based on pilot testing and results from previous studies with children (McCreery et al., 2010), four SNRs were used to obtain performance-intensity functions (0, +3, +6, +9). These levels were chosen to provide a range of SII values. To limit the length of the listening task and minimize the likelihood of changes in performance due to fatigue and/or decreased attention, each participant listened to two of four possible SNRs for each filter condition. For example, each participant would listen to either the +9/+3 dB SNRs or the +6/0 dB SNRs for each filter condition. Subsequent listeners within the same age group (children: 5–6 years, 7–8 years, 9–10 years, 11–12 years, and adults) completed the other SNRs for each filter condition. In all, each participant listened to two SNR conditions for each filter setting (7) for a total of 14 conditions. Participants were given one or two short breaks during the task depending on their age. The task typically took 90 min for children and 60 min for adults.

Correct/incorrect scoring of nonwords was completed online during the listening task. After the session, recordings of each participant were reviewed and scored to cross-check online scoring, as well as to analyze the phonemes in each response as correct or incorrect. Phonemes were judged to correspond to one of the phonemes in the stimulus set or were placed in a category for responses that were either unintelligible or not in the phonemes used to construct the nonwords for the current investigation. Confusion matrices were created for each subject and listening condition to allow for an analysis of specific consonant errors that may have contributed to differences in nonword recognition.

## III. RESULTS

### A. Nonword recognition

Prior to statistical analysis, nonword recognition scores were converted to Rationalized Arcsine Units (RAU) (Studebaker, 1985) to normalize variance across conditions. Because each child only listened to half of the SNR conditions for each filter condition, nonword recognition results represent combined results between two children within the same age group. Nonword recognition in the LP3 condition were consistently near 0% correct for all subjects and were excluded from subsequent statistical analyses due to lack of variance. To evaluate changes in nonword recognition as a function of age, a factorial repeated-measures analysis of variance (ANOVA) was completed with stimulus bandwidth and SNR as within-subjects factors and age-group (children: 5–6 years, 7–8 years, 9–10 years, 11–12 years, and adults) as a between-subjects factor. The main effect of age group was significant, $F(4,61) = 17.687$, $p < 0.001$, $\eta_p^2 = 0.537$, indicating that nonword recognition was significantly different across age groups. To evaluate the pattern of significant differences while controlling for Type I errors, *post hoc* comparisons were completed using Tukey's Honestly Significant Difference (HSD) with a minimum mean significant difference of 7.1 RAU. Adults had significantly higher

performance than the children in all four age groups. The mean differences between 5–6 year-olds and 7–8 year-olds (6.14 RAU) and between the 9–10 year-olds and 11–12 year-olds (0.41 RAU) were not significant. However, the 9–10 year-olds and 11–12 year-olds had significantly higher nonword recognition than the two younger groups. Based on this pattern and the lack of significant higher-order interactions involving age group, data for children are plotted by two age groups: younger children (ages 5 years: 0 months to 8 years: 11 months; $n = 56$) and older children, (ages 9 years: 0 months to 12 years: 11 months; $n = 60$). Data for adults are plotted seperately ($n = 19$). Mean nonword recognition for each age group and condition are shown in Figs. 2 and 3.

The main effect for stimulus bandwidth, $F$ (5,57) $= 354.709$, $p < 0.001$, $\eta_p^2 = 0.969$, was significant. Post hoc testing using Tukey's HSD with a minimum mean difference of 9.6 RAU revealed the highest nonword recognition scores in the full bandwidth conditions, and significant degradation in nonword recognition scores for each subsequent high- and low-pass filtering condition. The main effect of SNR was also significant, $F$ (5,57) $= 354.709$, $p < 0.001$, $\eta_p^2 = 0.969$, with the expected pattern of decreasing nonword recognition as SNR decreased with significant differences between all four SNR conditions on post hoc tests based on Tukey's HSD with a minimum mean difference of 4.8 RAU. The two-way interaction between stimulus bandwidth and SNR was significant, $F$ (15,915) $= 8.804$, $p < 0.001$, $\eta_p^2 = 0.126$, suggesting that the pattern of decreasing speech recognition for SNR differed across stimulus bandwidth conditions. As anticipated, degradation in performance with decreasing SNR increased with systematic decreases in stimulus bandwidth; however, this pattern was not observed for all bandwidth conditions. In general, decreases in nonword recognition for the same SNR were greater for low-pass listening conditions than for high-pass listening conditions. Post hoc tests using Tukey's HSD with a minimum mean significant difference of 8.6 RAU revealed different patterns of results for full bandwidth, high-pass, and low-pass conditions. For the full bandwidth condition, significant changes in nonword recognition were observed across all four SNRs. For the low-pass conditions, a similar pattern of significant differences was observed across all four SNRs until nonword recognition reached floor levels of performance. For high-pass conditions, performance differences between +6 and +9 dB SNR were not significant, but significant differences were observed between the +6, +3 and 0 dB SNR conditions. For children, EVT standard score was significantly correlated with mean nonword recognition score ($r = 0.25$, $p < 0.001$).

## B. Frequency-importance weights

Nonword recognition scores were used to derive frequency-importance weights for octave band frequencies for the adults, older children, and younger children. To obtain the importance of each octave band to the nonword
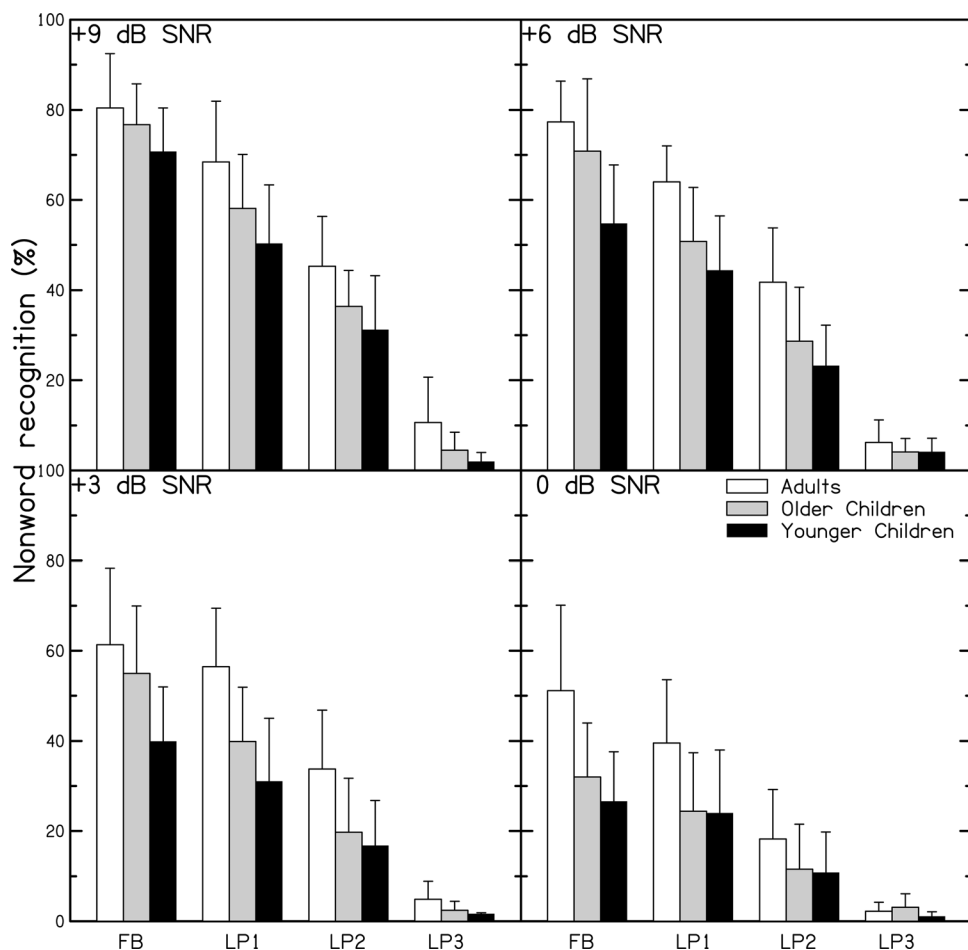


FIG. 2. Nonword recognition (% correct) for full bandwidth and low-pass conditions for adults (white), older children (gray) and younger children (black). Error bars are standard deviations. Each panel is SNR.
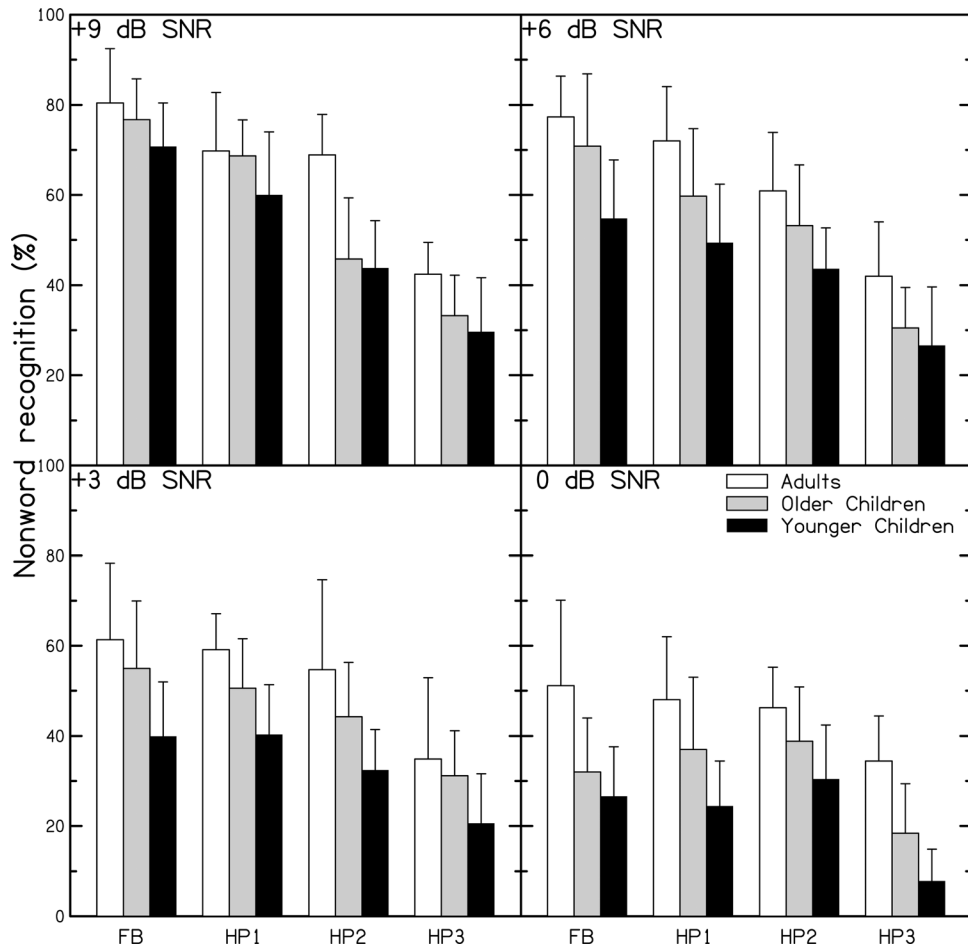
FIG. 3. Nonword recognition (% correct) for full bandwidth and high-pass conditions for adults (white), older children (gray) and younger children (black). Error bars are standard deviations. Each panel is SNR.

recognition score, the mean proportion of nonwords correct was calculated for each condition. The importance of each octave band was the amount of degradation in nonword recognition that was observed when that octave band was excluded. For example, the 8000 Hz importance weight was calculated as the mean difference between the full bandwidth condition for each subject and the low-pass condition without 8000 Hz (LP1). Derived frequency-importance functions are plotted in Fig. 4 with the nonword importance function from the current ANSI standard.
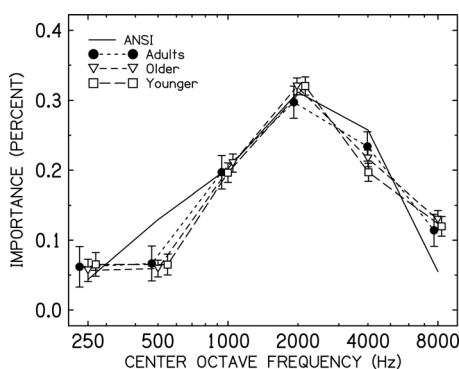


FIG. 4. Frequency-importance weight as a function of octave frequency band for the ANSI standard for SII for nonword syllables and for the current study groups (Adults—small dashed line; Older children—medium dashed line; Younger children—large dashed line). Error bars are standard deviations.

An mixed factorial analysis of variance with frequency-importance weight and SNR as within-subjects factors and age-group as a between-subjects factor revealed no significant differences across the three age groups, $F (10,315) = 1.088$, $p = 0.371$, $\eta_p^2 = 0.033$. Additionally, the frequency band by SNR interaction was not significant, $F (15,945) = 0.40$, $p = 0.115$, $\eta_p^2 = 0.023$, suggesting that the pattern of importance weights for each frequency band did not vary significantly by SNR.

## C. Transfer functions

The purpose of derived transfer functions is to allow estimation of speech recognition from the SII. In the current study, the transfer function for each age group was also calculated to examine the accuracy of SII predictions of speech recognition as a function of age. The transfer functions were derived as in previous studies (Scollie, 2008; Studebaker and Sherbecoe, 1991, 2002) using the equation

$$S = (1 - 10^{-SII/Q})^N, \tag{2}$$

where $S$ is the proportion correct speech recognition score, SII is the calculated SII for each condition and Q and N are fitting constants that define the slope and curvature of the transfer function, respectively. A nonlinear regression with SII as a predictor and nonword recognition as the outcome converged in 12 iterations revealed that the SII accounted

R. W. McCreery and P. G. Stelmachowicz: Audibility and speech recognition

TABLE III. Parameter estimates for non-linear regression used to estimate transfer functions for each age group.

| Age group | Parameter | Estimate | Std. error | 95% Confidence interval | |
| --- | --- | --- | --- | --- | --- |
| | | | | Lower | Upper |
| Adults | Q | 0.352 | 0.013 | 0.326 | 0.377 |
| | N | −1.830 | 0.036 | −1.900 | −1.760 |
| Older | Q | 0.450 | 0.011 | 0.428 | 0.471 |
| | N | −1.918 | 0.019 | −1.956 | −1.881 |
| Younger | Q | 0.570 | 0.015 | 0.539 | 0.600 |
| | N | −1.987 | 0.019 | −2.022 | −1.953 |

for 78% of the variance in adult nonword recognition scores with an RMS error of 4.1 RAU. The transfer functions for the older and younger children were fit using the same non-linear regression approach. For the older group of children, the solution converged in five iterations and accounted for 67.6% of the variance in nonword recognition with an RMS error of 11.1 RAU, whereas for the younger group the solution converged in six iterations and accounted for 65.9% of the variance in nonword recognition with an RMS error of 8.2 RAU. Regression coefficients for all three age groups are displayed in Table III. The predicted speech recognition scores as a function of SII for the adults, and the older and younger children are shown in Fig. 5.

## D. Consonant error patterns

Consonant error patterns were compared across age groups and filter conditions. In order to estimate the influence of these factors on consonant error patterns, proportion of reduction in error (PRE) was used (Reynolds, 1984). PRE analysis is a nonparametric statistical approach for categorical data that has been used previously to examine systematic variability in phoneme error patterns (Strange *et al.*, 2001). The PRE is based on the non-modal response for each consonant. An overall rate of non-modal responses for each consonant across conditions is calculated. This base error rate is compared to error rates for specific conditions or age groups to determine the reduction in error rate that occurs for that consonant for a specific group or condition. The difference
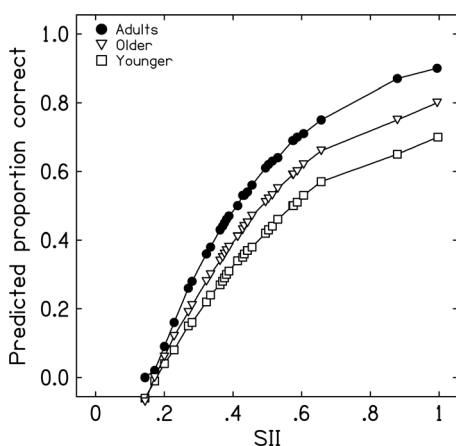


FIG. 5. Predicted proportion correct plotted across SII derived for data from the current study for different age groups (Adults—Filled circles; Older children—open triangles, Younger children—open squares).

between the base error rate and condition error rate are divided by the base error rate to estimate the PRE. As in previous studies (Strange *et al.*, 1998), a PRE of at least 0.10, reflecting a 10% reduction in error, was considered to reflect a systematic change in consonant recognition. Two patterns of consonant errors were examined. First, the effect of filter conditions on consonant error patterns was examined by comparing the error rate for each consonant across adjacent filter conditions to derive a PRE value for each frequency band. Similar to the procedure for deriving frequency-importance weights for each band, the PRE for a frequency band was calculated by taking the PRE between the two conditions where a specific band was removed by filtering. For example, the 8 kHz PRE was calculated by taking the PRE between the full bandwidth and LP1 condition where 8 kHz was filtered. Consonant error patterns that varied systematically across age groups were also identified. The target consonant represented the modal response for all conditions and groups with the exception of /h/, for which the modal response was /p/.

Table IV displays the consonant error rates for each filter condition and the PRE for each filter band. Errors in consonant recognition matched predictions based on the audibility of the frequency range of their acoustic cues. Recognition declined rapidly as the frequency bands corresponding to the acoustic cues of those consonants were removed by filtering. For example, fricative perception declined rapidly for each consecutive low-pass filtering condition, as would be expected for speech sounds comprised of high-frequency energy. All consonants showed significant PRE as mid and high frequency bands were removed. PRE for high-pass conditions revealed a significant PRE for affricates and nasals, as well as the voiceless fricatives /s/, /θ/ and the voiceless stop /t/. While /s/ and /t/ were primarily confused with their respective voiced cognates /z/ and /d/ in high-pass conditions, /θ/ showed the greatest number of confusions with /f/ and its voiced cognate /ð/. As observed in previous studies (Miller and Nicely, 1955), consonant error patterns for high-pass conditions were less predictable and systematic than error patterns for low-pass conditions, as the low frequencies contain acoustic cues for place, manner and voicing. Whereas the error patterns for low-pass conditions are predictable based on the spectrum of acoustic cues for specific consonants, the consonant error patterns for high-pass conditions are the result of the alteration of multiple acoustic cues within the same condition, leading to less consistent patterns across listeners.

Table V displays the PRE results across age group. Differences in consonant error patterns across age group were examined by using the adult or older child groups as the base error rate and calculating the PRE across age groups. Age-related differences in phoneme recognition helped to account for the observed age-dependency of nonword recognition. Younger children had higher consonant error rates than adults for all consonants except /ð/, /dʒ/, and /θ/. Older children had higher errors rates than adults for /g/, /k/, /p/, /tʃ/, /ʃ/, /m/, /n/, and /ŋ/. Younger children had higher consonant error rates than older children for /b/, /g/, /k/, /p/, /tʃ/, /f/, /s/, /ʃ/, /v/, /z/, /ʒ/, /m/, /n/,

| | Filter condition base error rate | | | | | | | PRE by band frequency (kHz) | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | FB[a] | LP1 | LP2 | LP3 | HP1 | HP2 | HP3 | 0.25 | 0.5 | 1 | 2 | 4 | 8 |
| p | 0.44 | 0.54 | 0.54 | 0.74 | 0.46 | 0.31 | 0.33 | 0.04 | **−0.33**[b] | 0.07 | **0.37** | 0.00 | **0.22** |
| b | 0.58 | 0.62 | 0.65 | 0.75 | 0.55 | 0.47 | 0.43 | −0.06 | **−0.14** | −0.09 | **0.16** | 0.03 | 0.07 |
| t | 0.05 | 0.08 | 0.65 | 0.78 | 0.05 | 0.07 | 0.12 | −0.03 | **0.40** | **0.85** | **0.19** | **7.43** | **0.59** |
| d | 0.14 | 0.18 | 0.59 | 0.81 | 0.14 | 0.12 | 0.12 | −0.02 | **−0.15** | 0.01 | **0.38** | **2.28** | **0.28** |
| k | 0.16 | 0.15 | 0.29 | 0.80 | 0.12 | 0.11 | 0.18 | −0.08 | 0.04 | **0.58** | **1.73** | **0.98** | 0.09 |
| g | 0.22 | 0.22 | 0.27 | 0.67 | 0.16 | 0.14 | 0.26 | −0.09 | **−0.13** | **0.87** | **1.47** | **0.24** | 0.01 |
| m | 0.40 | 0.42 | 0.51 | 0.58 | 0.47 | 0.45 | 0.48 | **0.17** | 0.05 | 0.08 | **0.13** | **0.21** | 0.06 |
| n | 0.31 | 0.32 | 0.42 | 0.70 | 0.44 | 0.24 | 0.28 | **0.45** | **−0.46** | **0.17** | **0.67** | **0.31** | 0.05 |
| ŋ | 0.52 | 0.55 | 0.51 | 0.78 | 0.58 | 0.63 | 0.68 | **0.12** | 0.08 | 0.09 | **0.53** | 0.07 | 0.06 |
| tʃ | 0.09 | 0.12 | 0.40 | 0.81 | 0.11 | 0.14 | 0.24 | **0.22** | **0.25** | **0.69** | **1.03** | **2.24** | **0.33** |
| dʒ | 0.09 | 0.12 | 0.45 | 0.73 | 0.12 | 0.18 | 0.19 | **0.33** | **0.53** | 0.06 | **0.63** | **2.62** | **0.43** |
| h | 0.68 | 0.75 | 0.80 | 0.81 | 0.68 | 0.57 | 0.64 | −0.01 | **−0.16** | **0.13** | 0.01 | 0.08 | 0.09 |
| f | 0.36 | 0.44 | 0.52 | 0.68 | 0.39 | 0.40 | 0.25 | 0.07 | 0.02 | **−0.36** | **0.31** | **0.17** | **0.22** |
| v | 0.37 | 0.42 | 0.47 | 0.67 | 0.36 | 0.30 | 0.27 | −0.02 | **−0.18** | **−0.10** | **0.43** | **0.11** | **0.15** |
| s | 0.12 | 0.21 | 0.89 | 0.80 | 0.10 | 0.16 | 0.17 | **0.11** | **0.52** | 0.07 | **0.10** | **3.31** | **0.78** |
| z | 0.12 | 0.24 | 0.84 | 0.73 | 0.09 | 0.10 | 0.12 | **0.22** | **0.13** | **0.15** | **−0.13** | **2.53** | **1.07** |
| θ | 0.51 | 0.69 | 0.74 | 0.82 | 0.55 | 0.70 | 0.68 | 0.07 | **0.28** | −0.03 | **0.11** | 0.07 | **0.35** |
| ð | 0.54 | 0.66 | 0.69 | 0.80 | 0.58 | 0.67 | 0.60 | 0.07 | **0.14** | −0.09 | **0.16** | 0.04 | **0.21** |
| ʃ | 0.18 | 0.25 | 0.33 | 0.76 | 0.16 | 0.18 | 0.17 | −0.01 | **0.11** | −0.04 | **1.31** | **0.32** | **0.39** |
| ʒ | 0.29 | 0.23 | 0.46 | 0.68 | 0.25 | 0.23 | 0.19 | −0.04 | −0.06 | **−0.18** | **0.49** | **1.03** | **−0.22** |

[a]FB = full bandwidth, LP = low-pass, HP = high-pass.
[b]PRE values in bold represent a difference that exceeds the 0.10 criterion.

and /ŋ/. Analysis of error patterns suggested that children frequently confused /k/ and /g/ with other stop consonants, whereas /p/ and /b/ were confused with stops and fricatives, particularly /f/ and /v/. The nasals /m/, /n/, and /ŋ/ showed an increasing pattern of recognition as a function of age group, as younger children were more likely to

confuse the nasals with other nasals than older children and adults.

## IV. DISCUSSION

The purpose of the current study was to evaluate predictions of speech recognition for children and adults based on the SII using nonword stimuli with limited contextual and linguistic cues. Overall, children had lower nonword recognition scores in noise than adults for the same amount of audibility as measured by the SII. Nonword recognition decreased predictably for all participants as the level of noise increased and spectral content became more limited. Despite significant differences between age groups in nonword recognition, the amount of degradation when octave bands were removed did not vary as a function of age as measured by differences in the frequency-importance weights across age groups. Age-related differences in nonword recognition are consistent with previous studies using a wide range of speech stimuli (Elliot, 1979; Nabalek and Robinson, 1982; Johnson, 2000; Scollie, 2008; Neuman *et al.*, 2010). However, the lack of differences between adults and children across conditions with varying bandwidth does not match the hypothesized effect or previous bandwidth effects observed for children.

TABLE V. Proportion of reduction in error and base error rates for each age group.

| | Young | Old | Adult | Y-A[a] | Y-O | O-A |
|---|---|---|---|---|---|---|
| p | 0.53 | 0.47 | 0.38 | **0.40**[b] | **0.14** | **0.23** |
| b | 0.65 | 0.54 | 0.50 | **0.31** | **0.20** | 0.09 |
| t | 0.27 | 0.26 | 0.24 | **0.12** | 0.06 | 0.06 |
| d | 0.32 | 0.29 | 0.27 | **0.20** | **0.10** | 0.09 |
| k | 0.30 | 0.24 | 0.19 | **0.59** | **0.25** | **0.28** |
| g | 0.32 | 0.25 | 0.20 | **0.59** | **0.27** | **0.26** |
| m | 0.52 | 0.46 | 0.36 | **0.45** | **0.13** | **0.28** |
| n | 0.43 | 0.38 | 0.30 | **0.46** | **0.14** | **0.27** |
| ŋ | 0.66 | 0.60 | 0.51 | **0.30** | **0.10** | **0.19** |
| tʃ | 0.32 | 0.29 | 0.24 | **0.33** | **0.12** | **0.20** |
| dʒ | 0.31 | 0.30 | 0.29 | 0.06 | 0.05 | 0.01 |
| h | 0.73 | 0.67 | 0.65 | **0.12** | 0.09 | 0.03 |
| f | 0.47 | 0.41 | 0.40 | **0.17** | **0.16** | 0.01 |
| v | 0.47 | 0.37 | 0.34 | **0.40** | **0.29** | 0.08 |
| s | 0.41 | 0.36 | 0.34 | **0.21** | **0.16** | 0.05 |
| z | 0.37 | 0.33 | 0.30 | **0.23** | **0.13** | 0.08 |
| θ | 0.69 | 0.68 | 0.64 | 0.06 | 0.01 | 0.05 |
| ð | 0.64 | 0.68 | 0.68 | −0.06 | −0.06 | 0.00 |
| ʃ | 0.35 | 0.30 | 0.25 | **0.39** | **0.17** | **0.19** |
| ʒ | 0.44 | 0.31 | 0.38 | **0.14** | **0.41** | **−0.19** |

[a]Y = Young children, O = Older children, A = Adults.
[b]PRE values in bold represent a difference that exceeds the 0.10 criterion.

## A. Predictions of speech recognition based on the SII

Within the age range of children in the current study, older children performed better than younger children for listening conditions with the same SII. Results from Scollie (2008) were consistent with the present findings, despite the use of different stimuli and different frequency-importance

weights to calculate SII values. The present findings suggest that while the SII is useful to quantify audibility of speech for children, conclusions about an individual child's speech recognition based on the SII are likely to overestimate performance unless age-specific data are used. Children had larger rms error for predictions of speech recognition based on the SII, reflecting greater variability for younger children. Scollie (2008) applied an age-related proficiency factor to limit the slope and asymptote of the transfer function to better predict speech recognition in children. The development of age-group specific SII predictions of speech recognition in this study is an empirical alternative to altering the adult function with a proficiency factor to better predict children's performance. Both age-based proficiency factors and predictions from age-specific data lead to better predictions than using adult data alone. However, variability within children of the same age, particularly for younger children where predictions are likely to be most valuable, may limit the applicability of either method to accurately estimate speech recognition for an individual listener.

While listening conditions with the same SII audibility would be expected to result in similar speech recognition outcomes, several conditions in the current study are not consistent with this prediction. As in previous work by Gustafson and Pittman (2011) where stimulus bandwidth and presentation level were both manipulated, nonword recognition in the current study varied across conditions with the same SII. This disparity is most obvious when comparing the LP1 and HP1 conditions for all four SNRs, which have essentially equivalent calculated SII. However, performance was significantly poorer in the LP1 conditions than HP1. The difference between observed and predicted performance could be related to the use of only female talkers with great high-frequency spectral content than male talkers used in previous studies. Future studies should seek to evaluate the factors that contribute to variability in speech recognition across conditions with the same calculated audibility.

Similar to results from previous studies of speech recognition in children (Elliot, 1979; Johnson, 2000; McCreery et al., 2010), nonword recognition was found to follow a predictable developmental pattern with adults achieving higher nonword recognition scores than both age groups of children with 9–12 year-old children performing better than 5–8 year-olds. Nonword stimuli were chosen for the current study to limit the influence of lexical knowledge and the ability to use the phonotactic characteristics of the stimuli on the recognition task. However, previous studies of nonword recognition in children have demonstrated that even when linguistic and phonotactic cues are constrained, nonword recognition tasks are strongly correlated with expressive vocabulary ability (Munson et al., 2005) and working memory (Gathercole, 2006). In the current study, expressive language scores as measured by the EVT-2 were significantly correlated with nonword recognition scores. Despite attempts to limit the influence of phonotactic probability on nonword repetition in the current study, the stimulus set was sufficiently large that stimuli with a wide range of phonotactic probabilities were included in the experiment. Age-related differences in nonword recognition were likely related to a combination of vocabulary ability, immediate memory skills and use of phonotactic probability.

## B. Frequency-importance weights as a function of age

Despite age-related differences in nonword recognition, the amount of degradation in nonword recognition observed across frequency bands did not vary as a function of age. This conclusion differs from the hypothesized effect of greater degradation for younger children when high-frequency bands are limited. Frequency-importance weights for children were expected to be more evenly distributed across bands resulting in greater importance at 4 kHz and 8 kHz than in adults, reflecting greater degradation in speech understanding when these bands were removed. In the current study, the distribution of importance weights also did not vary significantly across age groups or SNRs. These findings would seem to contradict a growing body of literature which suggests that children may be more negatively impacted than adults when bandwidth of speech is limited (Stelmachowicz et al., 2001; Pittman, 2008). Consistent with these previous bandwidth studies, however, fricatives with high-frequency spectral content, including /s/, /ʃ/, and /f/, showed developmental differences in error patterns with younger children showing greater degradation in identification as bandwidth decreased than older children and adults. Despite these differences in fricative perception, analysis of consonant patterns for each age group suggested that multiple phonemes including stops, nasals and fricatives contributed to age-related differences in nonword recognition. Consonant error patterns in the current study are similar to data from Nishi et al. (2010), where error patterns for younger children occurred across nearly all consonants and diminished for older children and adults. These results suggest that while fricative perception does vary as a function of the listener's age and the stimulus bandwidth, overall differences in nonword recognition are not related to one specific class of phonemes.

Additionally, several methodological aspects of the current study may have moderated the bandwidth effects for children observed in previous studies. Age-related differences in speech recognition due to limited bandwidth have been proposed to be related to children's development of linguistic knowledge and skills needed for top-down processing (Stelmachowicz et al., 2004). When acoustic-phonetic representations are not accessible due to noise or limited bandwidth, adults can rely on their understanding of language, context and phonotactic probability to help recognize degraded auditory stimuli. Because children are in the process of developing these skills, acoustic-phonetic factors such as broader stimulus bandwidth and higher SNR are needed to support decoding of the speech signal. The nonword stimuli used in the current experiment were controlled to limit the use of lexical and phonotactic cues for all listeners. Because all listeners had limited access to cues needed to support top-down processing, adults also relied on the acoustic-phonetic representation of the signal and developmental differences in frequency-importance were minimized. Adults are expected to perform better than children with a degraded signal for stimuli with redundant linguistic

cues, such as real words or sentences, where linguistic knowledge would be expected to support speech recognition.

Furthermore, nonword CVC stimuli used in the current investigation were balanced for the frequency of occurrence of initial and final consonants in order to promote comparability with nonword stimuli used in previous studies of frequency-importance functions with adults. Most previous bandwidth studies used stimuli with many fricatives or with stimuli that represent the frequency of occurrence for phonemes in English. Constraining the number of fricatives may have limited the observation of a bandwidth difference between low-pass conditions that would be necessary to observe differences in importance weights for high-frequency bands. Because /s/ is the third most frequently occurring phoneme in English (Denes, 1963), it is likely that a stimulus set with more fricatives or matched the occurance of phonemes in English may result in differences in frequency-importance in the bands where the acoustic energy for those phonemes occurs.

The frequency-importance weights derived for both children and adults in the current experiment are similar to those obtained with nonwords with phonemes occurring in equal frequencies that are used as the basis for the nonword importance weights in the ANSI standard. Two octave bands in the current study show different importance weights than the ANSI standard nonword weights. Specifically, the importance weight for the 8000 Hz band is higher than previous octave-band weights for nonwords. The difference is likely reflective of the use of female talkers with greater spectral content in that frequency band than the male talkers that were used in previous studies. These spectral differences are apparent in Fig. 1. Additionally, the 500 Hz band importance weight was significantly less than has been observed in previous studies, suggesting that subjects in the current study did not experience as much degradation in speech understanding when 500 Hz was removed from the stimulus as in previous research. Although the source of this difference is not clear, significant variability in speech recognition for adults in high-pass filtered listening conditions observed in the current study has also been reported in previously (Miller and Nicely, 1955). Because acoustic cues that signify place, manner and voicing occur in the low frequencies, error patterns are much less consistent and predictable for high-pass filtered conditions than low-pass filtered conditions.

## C. Limitations and future directions

Although the current study was the first to attempt to measure frequency-importance weights from children and adults using the same task, several limitations of the current study should be considered when comparing the results to previous studies and planning future research in this area. While efforts to reduce the task demands on children in the current study were necessary, constraining the number of conditions for each subject increased the variability of the results. Therefore, comparisons of the current results to previous studies where subjects listened to all conditions should be made cautiously. While the use of nonword CVCs allows the limitation of linguistic and phonotactic cues on speech recognition, these stimuli are likely to represent a worst-case

scenario for speech recognition, as even young children are able to use limited linguistic cues to support speech recognition (Boothroyd and Nittrouer, 1988).

Several unresolved questions could provide a basis for further study. The pattern of frequency-importance weights may be more likely to vary for stimuli with linguistic contextual cues, where adults and older children could more easily use cognitive and linguistic skills to support speech understanding. The role of linguistic cues in supporting speech recognition in children when bandwidth is limited also warrants further investigation. Because the differences between adults and children were not frequency-dependent, as evidenced by similar frequency-importance functions across age groups, potential mechanisms for age-related variability, such as differences in immediate memory, should be considered. Until further studies are conducted, estimates of speech recognition for children based on the SII should be viewed as likely to overestimate performance.

## D. Conclusion

The aim of the current study was to evaluate predictions of speech recognition for children and adults based on audibility as measured by the SII. Children between 5 and 12 years of age with normal hearing had poorer nonword recognition for listening conditions with the same amount of audibility compared to the performance of adults on the same task. However, contrary to previous studies, children did not experience greater degradation in speech recognition than adults when high-frequency bandwidth was limited. Adults and children both performed more poorly for band-limited conditions with stimuli with limited linguistic cues. This finding supports the need for maximizing high-frequency audibility in conditions where context is limited, particularly for young children who are developing linguistic knowledge and improving efficiency of related cognitive processes. The SII provides an estimate of audibility but use of the SII to predict speech recognition outcomes in children should take into account the potential variability both between adults and children and within children of the same age that were observed in the current investigation.

ANSI (**1997**). ANSI S3.5–1997, American National Standard Methods for Calculation of the Speech Intelligibility Index (American National Standards Institute, New York).

Bagatto, M., Scollie, S. D., Hyde, M., and Seewald, R. (**2010**). "Protocol for the provision of amplification within the ontario infant hearing program," Int. J. of Aud. **49**(Suppl. 1), S70–S79.

Bankson, N. W., and Bernthal, J. E. (**1990**). *Quick Screen of Phonology* (Special Press, Inc., San Antonio, TX), pp. 1–30.

Bell, T. S., Dirks, D. D., and Trine, T. D. (**1992**). "Frequency-importance functions for words in high- and low-context sentences," J. Speech Hear. Res. **35**, 950–959.

Boothroyd, A., and Nittrouer, S. (**1988**). "Mathematical treatment of context effects in phoneme and word recognition," J. Acoust. Soc. Am. **84**, 101–114.

Ching, T. Y., Dillon, H., and Byrne, D. (**1998**). "Speech recognition of hearing-impaired listeners: Predictions from audibility and the limited role of high-frequency amplification," J. Acoust. Soc. Am. **103**, 1128–1140.

Ching, T. Y., Dillon, H., Katsch, R., and Byrne, D. (**2001**). "Maximizing effective audibility in hearing aid fitting," Ear Hear **22**, 212–224.

Denes, P. B. (**1963**). "On the statistics of spoken English," J. Acoust. Soc. Am. **35**, 892–904.

Dubno, J. R., Dirks, D. D., and Schaefer, A. B. (**1989**). "Stop-consonant recognition for normal-hearing listeners and listeners with high-frequency hearing loss. II: Articulation index predictions," J. Acoust. Soc. Am. **85**, 355–364.

Duggirala, V., Studebaker, G. A., Pavlovic, C. V., and Sherbecoe, R. L. (**1988**). "Frequency importance functions for a feature recognition test material," J. Acoust. Soc. Am. **83**, 2372–2382.

Edwards, J., Beckman, M. E., and Munson, B. (**2004**). "The interaction between vocabulary size and phonotactic probability effects on children's production accuracy and fluency in nonword repetition," J. Speech Hear. Res. **47**, 421–436.

Eisenberg, L. S., Shannon, R. V., Martinez, A. S., Wygonski, J., and Boothroyd, A. (**2000**). "Speech recognition with reduced spectral cues as a function of age," J. Acoust. Soc. Am. **107**, 2704–2710.

Elliott, L. L. (**1979**). "Performance of children aged 9 to 17 years on a test of speech intelligibility in noise using sentence material with controlled word predictability," J. Acoust. Soc. Am. **66**, 651–653.

Fletcher, H., and Galt, R. H. (**1950**). "The perception of speech and its relation to telephony," J. Acoust. Soc. Am. **22**, 89–151.

French, N. R., and Steinberg, J. C. (**1947**). "Factors governing the intelligibilty of speech sounds," J. Acoust. Soc. Am. **19**, 90–119.

Gathercole, S. E. (**2006**). "Nonword repetition and word learning: The nature of the relationship," App. Psycholinguistics. **27**, 513–543.

Gustafson, S. J., and Pittman, A. L. (**2011**). "Sentence perception in listening conditions having similar speech intelligibility indices," Int. J. Audiol. **50**, 34–40.

Hnath-Chisolm, T. E., Laipply, E., and Boothroyd, A. (**1998**). "Age-related changes on a children's test of sensory-level speech perception capacity," J. Speech Hear. Res. **41**, 94–106.

Humes, L. E. (**2002**). "Factors underlying the speech-recognition performance of elderly hearing-aid wearers," J. Acoust. Soc. Am. **112**, 1112–1132.

Johnson, C. E. (**2000**). "Children's phoneme identification in reverberation and noise," J. Speech Hear. Res. **43**, 144–157.

Kortekaas, R. W., and Stelmachowicz, P. G. (**2000**). "Bandwidth effects on children's perception of the inflectional morpheme /s/: Acoustical measurements, auditory detection, and clarity rating," J. Speech Hear. Res. **43**, 645–660.

McCreery, R., Ito, R., Spratford, M., Lewis, D., Hoover, B., and Stelmachowicz, P. G. (**2010**). "Performance-intensity functions for normal-hearing adults and children using computer-aided speech perception assessment," Ear Hear **31**, 95–101.

Miller, G. A., and Nicely, P. E. (**1955**). "An analysis of perceptual confusions among some English consonants," J. Acoust. Soc. Am. **27**, 338–352.

Mlot, S., Buss, E., and Hall, J. W., (**2010**). "Spectral integration and bandwidth effects on speech recognition in school-aged children and adults," Ear Hear **31**, 56–62.

Moore, B. C. J., Stone, M. a, Füllgrabe, C., Glasberg, B. R., and Puria, S. (**2008**). "Spectro-temporal characteristics of speech at high frequencies, and the potential for restoration of audibility to people with mild-to-moderate hearing loss," Ear Hear **29**, 907–922.

Munson, B., Kurtz, B. A., and Windsor, J. (**2005**). "The influence of vocabulary size, phonotactic probability, and wordlikeness on nonword repeti-

tions of children with and without specific language impairment," J. Speech Hear. Res. **48**, 1033–1047.

Nabelek, A. K., and Robinson, P. K. (**1982**). "Monaural and binaural speech perception in reverberation for listeners of various ages," J. Acoust. Soc. Am. **71**, 1242–1248.

Neuman, A. C., Wroblewski, M., Hajicek, J., and Rubinstein, A. (**2010**). "Combined effects of noise and reverberation on speech recognition performance of normal-hearing children and adults," Ear Hear **31**, 336–344.

Nishi, K., Lewis, D. E., Hoover, B. M., Choi, S., and Stelmachowicz, P. G. (**2010**). "Children's recognition of American English consonants in noise," J. Acoust. Soc. Am. **127**, 3177–3188.

Nittrouer, S., and Boothroyd, A. (**1990**). "Context effects in phoneme and word recognition by young children and older adults," J. Acoust. Soc. Am. **87**, 2705–2715.

Pavlovic, C. V. (**1987**). "Derivation of primary parameters and procedures for use in speech intelligibility predictions," J. Acoust. Soc. Am. **82**, 413–422.

Pittman, A. L. (**2008**). "Short-term word-learning rate in children with normal hearing and children with hearing loss in limited and extended high-frequency bandwidths," J. Speech Hear. Res. **51**, 785–797.

Reynolds, H. T. (**1984**) *Analysis of Nominal Data*, 2nd ed. (Sage, Beverly Hills, CA), pp. 22–88.

Scollie, S. D. (**2008**). "Children's speech recognition scores: The speech intelligibility index and proficiency factors for age and hearing level," Ear Hear **29**, 543–556.

Seewald, R., Moodie, S., Scollie, S., and Bagatto, M. (**2005**). "The Desired Sensation Level method for pediatric hearing instrument fitting: Historical perspective and current issues," Trends Amplif. **9**, 145–157.

Stelmachowicz, P. G., Hoover, B. M., Lewis, D. E., Kortekaas, R. W., and Pittman, A. L. (**2000**). "The relation between stimulus context, speech audibility, and perception for normal-hearing and hearing-impaired children," J. Speech Hear. Res. **43**, 902–914.

Stelmachowicz, P. G., Pittman, A. L., Hoover, B. M., and Lewis, D. E. (**2001**). "Effect of stimulus bandwidth on the perception of /s/ in normal- and hearing-impaired children and adults," J. Acoust. Soc. Am. **110**, 2183–2190.

Stelmachowicz, P. G., Pittman, A. L., Hoover, B. M., Lewis, D. E., and Moeller, M. P. (**2004**). "The importance of high-frequency audibility in the speech and language development of children with hearing loss," Arch. Otolaryngol. Head Neck Surg. **130**, 556–562.

Storkel, H. L., and Hoover, J. R. (**2010**). "An online calculator to compute phonotactic probability and neighborhood density on the basis of child corpora of spoken American English," Behav. Res. Methods **42**, 497–506.

Strange, W., Akahane-Yamada, R., Kubo, R., Trent, S. A., Nishi, K., and Jenkins, J. J. (**1998**). "Perceptual assimilation of American English vowels by Japanese listeners," J. Phonetics **26**, 311–344.

Strange, W., Akahane-Yamada, R., Kubo, R., Trent, S. A., and Nishi, K. (**2001**). "Effects of consonantal context on perceptual assimilation of American English vowels by Japanese listeners," J. Acoust. Soc. Am. **109**, 1691–1704.

Studebaker, G. A. (**1985**). "A 'rationalized' arcsine transform," J. Speech Hear. Res. **28**, 455–462.

Studebaker, G. A., Pavlovic, C. V., and Sherbecoe, R. L. (**1987**). "A frequency importance function for continuous discourse," J. Acoust. Soc. Am. **81**, 1130–1138.

Studebaker, G. A., and Sherbecoe, R. L. (**1991**). "Frequency-importance and transfer functions for recorded CID W-22 word lists," J. Speech Hear. Res. **34**, 427–438.

Studebaker, G. A., and Sherbecoe, R. L. (**2002**). "Intensity-importance functions for bandlimited monosyllabic words," J. Acoust. Soc. Am. **111**, 1422–1436.

Studebaker, G. A., Sherbecoe, R. L., and Gilmore, C. (**1993**). "Frequency-importance and transfer functions for the Auditec of St. Louis recordings of the NU-6 word test," J. Speech Hear. Res. **36**, 799–807.

Williams, K. T. (**2007**). *Expressive Vocabulary Test*, 2nd ed. (Pearson, Minneapolis, MN), pp. 1–225.