Sentence Perception of Children and Adults in a Series of Listening Conditions having Equivalent Speech Intelligibility Indices

Samantha Gustafson and Andrea Pittman

Department of Speech and Hearing Sciences, Arizona State University

Abstract

The purpose of this study was to examine directly the interaction between sensation level and bandwidth under conditions of equivalent SII. Participants were 20 adults between the ages of 19-47 years and 20 8-year-old children. All participants had normal hearing. Participants were asked to repeat four-word sentences that were filtered in four low-pass conditions (0.8, 1.25, 2.5, 8.0 kHz) having equal SII values of 0.40 - 0.41. The sentences were either high- or low-predictability in content and presented in noise shaped to produce masked hearing thresholds of 40 dB SPL at octave frequencies. The sentence perception data were subjected to a repeated-measures ANOVA with group and sentence type as the between-subject factors and low-pass filter condition as the within-subject factors. Results indicated a significant main effect for group, sentence type, and low-pass filter condition. Additionally, no significant group x condition, group x sentence type, or condition x sentence type interactions were revealed. Post-hoc analysis of the significant main effect of low-pass filter condition revealed equal performance at 1.25, 2.5, and 8.0 kHz filter conditions for both children and adults when repeating high-predictability sentences. These results suggest that for familiar speech, performance can be reasonably predicted by the SII for bandwidths $\geq$1.25 kHz. However, the same pattern of performance was not observed for the low-predictability sentences suggesting that the SII may be inappropriate for predicting the perception of children listening to unfamiliar speech.

Introduction

Children with hearing impairment (HI) have been shown to produce and recognize speech more poorly than their normal-hearing (NH) peers. Research has shown that a child's ability to perceive speech directly affects his/her ability to produce speech (Boothroyd, 1984; Elfenbein, 1994; Moeller, 2007; Oller, Eilers, Bull, & Carney, 1985).

To predict and compare the performance of listeners with NH and HI, a standardized procedure called the Speech Intelligibility Index (SII) was developed (ANSI, 1997) and is summarized in the following formula:

$$SII = \frac{1}{30}\sum_{i=1}^{18} SL_i W_i$$

The SII is the sum of the suprathreshold levels in discrete frequency bands multiplied by the degree to which each frequency band contributes to speech intelligibility. The resulting value is a number between 0 and 1. A signal with an SII of 0 would be inaudible to the listener and therefore performance would be 0% whereas a signal with an SII of 1 is completely audible and would predict a performance of 100%. The relative importance of each frequency band was derived by using a subjective speech task in which adults with normal hearing were asked to rate the clarity and intelligibility of the signal (Studebaker, 1987). One consequence of this approach is the assumption that the sensation level in one frequency band may be traded with, or redistributed to, another frequency band as long as the overall contribution across bands remains constant.

A number of studies have used the SII to equate audibility across listeners without controlling for the distribution of sensation level across frequencies (Stelmacowicz et al., 2000; Ching et al., 1998). The results of these previous studies showed large variability in performance at equal audibility levels. The purpose of this study was to examine directly the interaction between sensation level and bandwidth under conditions of equivalent SII.

Methods

Participants:

Participants were 20 adults between the ages of 19-47 years and 20 8-year-old children. All participants had normal hearing.

Stimuli:

Stimuli were 4-word sentences that were either high- or low-predictability in content. The high-predictability (HP) sentences were both grammatically and semantically correct (e.g. "Pick up this room"). Low-predictability (LP) sentences were grammatically correct, but semantically meaningless (e.g. "Jokes sleep on fields"). HP sentences represent familiar communication whereas LP sentences represent information which is unfamiliar to the listener. The stimuli were produced by a female talker with a standard American English dialect. The recordings were digitized at a sampling rate of 22.05 kHz.

The stimuli were frequency shaped to produce four low-pass filter conditions (0.8, 1.25, 2.5, and 8.0 Hz). The amplitude and cut-off frequencies were carefully adjusted to provide equal audibility (SII of 0.40 – 0.41) for each condition (Fig 1). Speech intelligibility was calculated using the frequency importance function for short passages (ANSI, 1997).

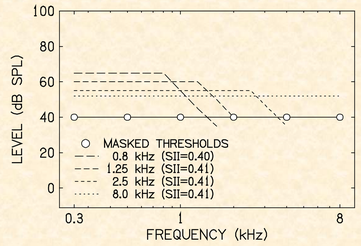

Figure 1. Stimulus level as a function of frequency for the four low-pass filter conditions and the masked hearing thresholds.

Procedure:

- Otoscopy and tympanometry were performed to confirm normal middle-ear function.
- The Peabody Picture Vocabulary Test was administered to each child to determine receptive vocabulary scores.
- Testing was conducted in one ear only. Masked thresholds were obtained using a broadband noise that was frequency shaped to produce thresholds of 40 dB SPL at octave frequencies from 0.25 through 8.0 kHz.
- Perception of both the HP and LP sentences in each of the four low-pass filter conditions was measured.

Results

The sentence perception data were subjected to a repeated measures analysis of variance (ANOVA) with group (adults, children) and sentence type (HP, LP) as the between subjects factors, and low-pass filter condition (0.8, 1.25, 2.5, 8.0 kHz) as the within subjects factor. Post-hoc analyses (pairwise comparisons with a Bonferroni adjustment for multiple comparisons) were conducted to identify significant differences between the low-pass filter conditions. Significant main effects for group and sentence type and filter condition were revealed as well as an interaction between low-pass filter condition and sentence type.

Figure 2 shows average performance as a function of low-pass filter condition collapsed across sentence type. A significant main effect of group was revealed indicating that on average, the children's performance was 20-30% lower than that of the adults. No interaction between the low-pass filter condition and group was revealed ($F(3,114) = 1.595$; $p=0.194$) indicating that, although the children's performance was poorer, both groups responded similarly to each low-pass filter condition. That is, performance increased as bandwidth increased until the broadest low-pass filter condition in which performance decreased in both groups.

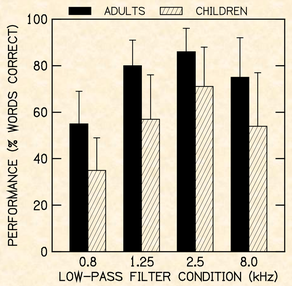

Figure 2: Average performance of adults and children in each low-pass filter condition collapsed across sentence type.

Results (con't)

Figure 3 shows average performance as a function of low-pass filter condition for the adults (upper panel) and the children (lower panel). Recall that audibility was equated across the four low-pass filter conditions. Conditions in which performance was statistically equivalent are indicated by horizontal lines. For the adults listening to HP sentences, the 0.8 kHz low-pass filter condition was statistically poorer than all other low-pass filter conditions. Performance increased significantly in the 1.25 kHz low-pass filter condition and remained constant for the 2.5 and 8.0 kHz low-pass filter conditions. The same improvement in performance was observed for the LP sentences. However, performance in the widest low-pass filter condition (8.0 kHz) decreased significantly. Like the adults, the children listening to HP sentences demonstrated a significant increase in performance for the low-pass filter conditions >0.8 kHz and performance remained constant. The children's performance with the LP sentences however, increased significantly up to the 2.5 kHz low-pass filter condition but decreased in the widest low-pass filter condition.

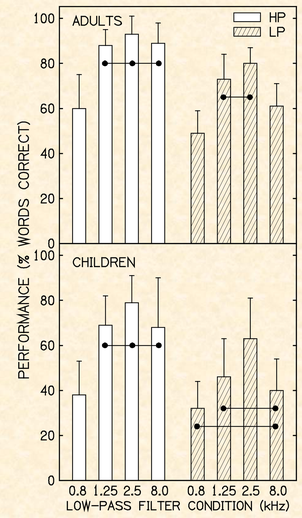

Figure 3. Average performance for adults and children in each low-pass filter condition for high- and low-predictability sentences.

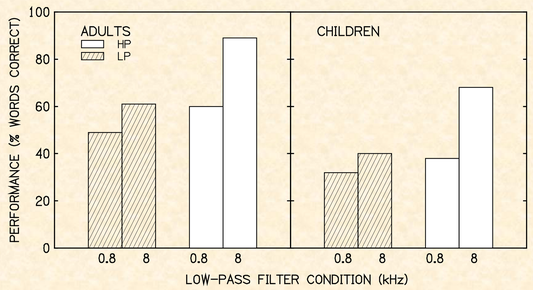

Figure 4. Average performance for adults and children in the narrowest and widest low-pass filter condition for each sentence type.

To determine the impact of bandwidth on the listener's ability to take advantage of the context within the sentence, performance in the narrowest and widest low-pass filter conditions for both the high- and low-predictability sentences were compared. Figure 4 shows the results for the adults (left panel) and the children (right panel). Performance for the LP sentences is shown on the left and for the HP sentences on the right. Similar performance was observed for the LP sentences in both filter conditions. However, performance increased significantly with the increased bandwidth for the HP sentences. These results suggest that although sensation level was similarly low for both the HP and LP sentences, the listeners were better able to take advantage of the contextual cues in the widest bandwidth condition only. Consequently, a listener may benefit more from a signal with a wide bandwidth at a lower level than from a signal with a narrow bandwidth presented at a higher level, even if the audibility was the same for each listening condition.

Conclusion

The present study was designed to explore the relationship between bandwidth and sensation level as it applies to the Speech Intelligibility Index. In this study the importance functions specified in the ANSI standard (1997) were used to hold audibility constant across a number of conditions. To do so, bandwidth and sensation level were systematically adjusted. Listeners were asked to repeat both low- and high-predictability sentences. It was expected that performance would be equal in each of the low-pass filter conditions. However, results showed that performance varied significantly across conditions in both children and adults. These results suggest an interaction between bandwidth and sensation level such that audibility cannot be effectively redistributed across the frequency domain. Therefore, similar performance cannot be expected based on SII calculations alone.

References

American National Standards Institute. Methods for Calculation of the Speech Intelligibility Index (ANSI S3.5-1997). New York: 1997.

Boothroyd, A. "Auditory perception of speech contrasts by subjects with sensorineural hearing loss." J.Speech Hear.Res. 27.1 (1984): 134-44.

Ching, T. Y., H. Dillon, and D. Byrne. "Speech recognition of hearing-impaired listeners: predictions from audibility and the limited role of high-frequency amplification." J Acoust.Soc.Am 103.2 (1998): 1128-40.

Elfenbein, J. L., M. A. Hardin-Jones, and J. M. Davis. "Oral communication skills of children who are hard of hearing." J.Speech Hear.Res. 37.1 (1994): 216-26.

Moeller, M. P., et al. "Vocalizations of infants with hearing loss compared with infants with normal hearing: Part I--phonetic development." Ear Hear. 28.5 (2007): 605-27.

Oller, D. K., et al. "Prespeech vocalizations of a deaf infant: a comparison with normal metaphonological development." J.Speech Hear.Res. 28.1 (1985): 47-63.

Stelmachowicz, P. G., et al. "The relation between stimulus context, speech audibility, and perception for normal-hearing and hearing-impaired children." J.Speech Lang Hear.Res. 43.4 (2000): 902-14.

Studebaker, G. A., et al. "Age-related changes in monosyllabic word recognition performance when audibility is held constant." J.Am.Acad.Audiol. 8.3 (1997): 150-62.