Consonant Enhancement Effects on Speech Recognition of Hearing-Impaired Children

Lynn Z. Smith*
Harry Levitt'

Abstract

Differences in gain (enhancement, in dB) required to optimize the consonant/vowel intensity ratio in nonsense syllables were determined for stops and fricatives, both voiced and voiceless, in 12 children with congenital moderate to severe sensorineural hearing loss. The test stimuli were vowel/consonant nonsense syllables with various levels of enhancement ranging from 0 dB (for the unprocessed stimulus) to 24 dB of gain, in steps of 3 or 6 dB. Results showed that significant improvements in consonant recognition can be obtained with individualized adjustment of consonant amplitude for children as young as 5 years of age.

Key Words: Children, hearing impaired, speech perception

Abbreviations: CE = consonant enhancement, CVR = consonant/vowel intensity ratio, NST = Nonsense Syllable Test

Understanding speech is difficult for individuals with sensorineural hearing loss. Comprehending what has been said can be like fitting pieces of a puzzle. Parts of the spoken message are often missing or distorted. The hearing-impaired listener must sort through the input using contextual information and language experience to arrive at meaning. Hearing-impaired children have more difficulty sorting information than adults because they have limited experience with language and a weaker base of world knowledge upon which to build contextual cues. Additionally, children may process auditory information differently than adults. However, both adults and children with sensorineural hearing impairment must deal with an incoming speech signal comprising features of mixed degrees of clarity.

In an attempt to determine what features of speech are difficult for hearing-impaired listeners, Picheny et al (1986) found that talkers use different strategies when speaking to the hearing impaired. They observed that speakers spontaneously adapt their speech when talking to individuals with hearing impairment. This type of “clear” speech is more easily understandable to hearing-impaired listeners than “conversational” speech or speech that has not been adapted for the benefit of the hearing impaired. Picheny et al (1986) analyzed both types of speech to determine what characteristics had been changed in clear speech. In general, three changes occurred: (a) the duration of the consonants and vowels increased, (b) the phonemes were articulated as if the words were in isolation, and (c) the consonant/vowel intensity ratio (CVR) increased. Very little change in formant frequencies was observed from conversational to clear speech.

Efforts to use Picheny et al’s findings have been proposed in the form of speech enhancement using digital technology (Montgomery and Edge, 1988). Whereas automatic processing of speech to alter articulatory patterns is unlikely to be accomplished in a wearable hearing aid in the near future, automatic adjustment of the CVR is a viable possibility (Preves, 1994). Speech characteristics that are weak, distorted, or imperceptible to hearing-impaired listeners may be enhanced through hearing aid circuitry that alters the gain for selected portions of the speech signal.

Whereas modification of CVR has been shown to produce improved consonant recognition with hearing-impaired adults, it is not yet known whether this technique will yield simi-
lar benefits for hearing-impaired children. (Revoile et al., 1986; Gordon-Salant, 1987; Montgomery and Edge, 1988; Freyman et al., 1991; Kennedy et al., 1998). There is evidence of increase in auditory abilities with age (Siegenthaler, 1969; Abromovitz, 1971; Neuman and Hochberg, 1983; Irwin et al., 1986; Allen et al., 1989; Werner and Rubel, 1992). As a consequence, hearing-impaired children may have amplification needs that require developmental considerations.

The present study focused on the effects of CVR alterations on consonant identification by hearing-impaired children. It sought to determine the amount of consonant enhancement (CE) needed to maximize consonant recognition. CE means the gain, in dB, applied to the consonant so as to increase the CVR. The optimal consonant enhancement (for maximizing consonant recognition) was determined for children for the two most common consonant types (stops and fricatives, both voiced and voiceless) in the syllable-final position, since this is the condition that is most difficult for consonant perception by hearing-impaired listeners.

METHOD

Subjects

Twelve congenitally hearing-impaired children with moderately to severely and gradually sloping sensorineural hearing loss acted as subjects. Moderately to severely hearing-impaired listeners were chosen for this study since they have been shown to benefit from CE (Gordon-Salant, 1987; Montgomery and Edge, 1988). One group ranged in age from 5 to 6 years and the other group was aged from 8 to 9 years. The younger group attended a preschool for the hearing impaired where emphasis was placed on the development of aural/oral communication skills. The older group attended mainstream programs in public schools and used oral communication. None of the subjects had any other handicapping conditions. All of the 12 subjects were experienced hearing aid users.

Figures 1 and 2 show composite audiograms of the mean threshold levels and ranges for the better ear (test ear) for the younger and older subjects, respectively. Because of constraints of age, audiometric configuration, and parental consent, it was not possible to have equal numbers of boys and girls in each group. Previous research on CVR adjustment has not shown any between-gender differences, and it is unlikely that different results would have been obtained had equal representation of genders been possible.

Test Stimuli

Test items from the Nonsense Syllable Test (NST) developed by Levitt and Resnick (1978) were used as stimuli. An adult male acted as the
speaker. The vowel/consonant (VC) syllables of interest were digitized using a 20 KHz sampling rate and stored in a PC-AT type computer using a 12-bit analog-to-digital converter (Data Translation 2801-A data acquisition board). The board was also used to provide digital-to-analog conversion for playing back the processed stimuli. Antialiasing and anti-imaging filters with a nominal cut-off frequency of 7.5 kHz were used. This cut-off frequency was conservatively below the theoretical maximum of 10 kHz to guard against spurious effects that might result from the finite slope of the filter skirts.

The frication noise portion of fricatives and the burst portion of stops in each VC of the digitized speech signal were isolated using a waveform editor. The precise start and stop times for each fricative or burst and vowel in each VC was determined using methodology similar to that described in Dubno and Levitt (1981). The unmodified (natural) speech was defined as having a CE of 0 dB. Test stimuli were then generated with CEs ranging from 0 dB (baseline) to a maximum of +24 dB in steps of 3 dB.

Two sets of consonants were used: (1) voiceless and voiced fricatives: /s, f, θ, z, v, ʎ/; and (2) voiceless and voiced stops: /p, t, k, b, d, g/. Each consonant followed an /a/ vowel in a VC format.

Test Procedure

Most comfortable loudness (MCL) and uncomfortable loudness (UCL) were obtained for each child. The nonsense syllables were then presented with the vowel at MCL. Those children whose dynamic ranges could accommodate the full range of CE without loudness discomfort were presented with five levels of CE, from 0 dB to 24 dB in 6-dB steps. For those children with narrower dynamic ranges, five levels of enhancement were used in roughly equal steps (to the nearest 3 dB) from 0 dB to the maximum intensity possible without loudness discomfort. Since the minimum step size for consonant enhancement was 3 dB, children with dynamic ranges of less than 12 dB received fewer levels of CE so as not to exceed their UCL levels. Dynamic ranges were determined separately for each consonant type.

Once the dynamic ranges and allowable levels of CE for each child were determined, the test stimuli were randomized and presented to the children in random order. Twelve replications of each consonant at each level of enhancement were presented to each child.

RESULTS

This investigation posed three questions:

1. What is the effect of CE on consonant recognition in children?
2. What is the effect of age on consonant recognition as a function of CE?
3. If CE produces an improvement in consonant recognition, what level of enhancement will maximize consonant recognition?

In order to answer the first question, the subjects' responses to all VCs at all levels of enhancement were analyzed. Percent correct scores for each subject and consonant type were plotted as a function of CE. These are referred to as CE functions.

The CE functions were fitted to the data using the method of orthogonal polynomials (Bennet and Franklin 1954; Levitt and Rabiner, 1971), taking into account the binomial variability of percentage scores. This was done by transforming the percent correct scores to arcsine units ($Y = 2 \arcsin (\sqrt{p/100})$, where $p$ equals observed score in percent) prior to fitting the curve. The highest point on the curve, $Y_{\text{max}}$, identifies the maximum score in arcsine units. This score, transformed back to percent correct, is referred to as Pmax. The level of CE in decibels corresponding to Pmax is defined as CEmax. The test score in percent for no enhancement (i.e., CE = 0 dB) is defined as Po, or Yo when arcsine units are used.

Several different types of CE functions were obtained. These functions were classified as no change, rising/peak, rising/no peak, or falling, as shown in Figure 3. The frequencies of occurrence of CE functions in each classification are shown in Figure 4 for each consonant type. The largest number of CE functions fall into the no change category. Of these, the most frequent case was for the voiced fricative /v/. Enhancement of the voiceless alveolar stop /t/ and the voiced labial stop /b/ also yielded large numbers of no change CE functions.

The categories of greatest interest are rising/peak and rising/no peak, since these represent the cases for which there is improved consonant recognition. The rising/peak and the rising/no peak CE functions were obtained most frequently with the voiceless glottal stop /k/ and the voiceless fricative /s/. The voiced alveolar stop /d/ also had a relatively large number of rising/no peak CE functions. The last category,
falling configurations, represents cases in which CE actually produced a decrement in consonant recognition. The voiced fricative /v/, the voiceless labial dental fricative /f/, and the voiced labial stop /b/ had the largest number of falling CE functions.

**Analysis of Consonant Recognition**

A five-factor, repeated-measures analysis of variance was performed on the consonant recognition scores. The factors were cognate pair (C), six levels, (/p/b, /t/d/, /k/g/, /f/v/, /θ/Ω, /s/z/); voicing (V), two levels (voiceless, voiced); enhancement (E), two levels (unenhanced, Po; maximum enhancement, Pmax); age group (A), two levels (younger, older); subjects (S), six per age group.

The measures were repeated over each subject. Note that, for obvious reasons, different subjects were used in each age group. This had the effect of lowering the denominator degrees of freedom in the F-tests for V, EA, and VEA from 11 (for 12 subjects) to 10 (for two groups of six subjects). Similarly, the denominator degrees of freedom in the F-tests for C, CV, and CE are reduced from 55 to 50. Note also that the factors place and manner were combined to form the joint factor cognate pair, which is orthogonal to voicing (i.e., voicing could thus be treated as a separate factor). On the other hand, the three places of articulation for the stops, bilabial (/p/b), alveolar (/t/d), and velar (/k/g), are not the same as the three places of articulation for the fricatives, labiodental (/v/), dental (/θ/Ω), and alveolar (/s/z); thus, place and manner are not orthogonal variables. It is for this reason that place and manner are treated as a single, joint factor rather than two independent factors.

Since the data are in the form of percentages, the arcsine transformation was used to stabilize the error variance. The data were transformed back to percentages on completion of the statistical tests.

The results of the analysis of variance are shown in Table 1. Of the four main effects, cognate pair (C) and enhancement (E) were found to be highly significant (p < .001). The main
effects voicing (V) and age group (A) were not statistically significant, although several interactions involving A and E were statistically significant. Significance levels for these interactions ranged from $p < .1$ to $p < .005$. The table shows only those interactions that were statistically significant. The significant main effects and interactions are summarized in Figure 5. The data are represented by histograms. The height of each bar corresponds to the percent correct score. For those main effects or interactions contrasting enhanced versus unenhanced conditions, solid bars are used to represent the unenhanced condition (Po) while open bars are used to represent the enhanced condition with the highest score (Pmax). The open bars are superimposed on the solid bars so that the portion of the open bar above the solid bar represents the increase in percent correct in going from Po to Pmax.

The main effects of cognate pair (C) and enhancement (E) are shown in the top third of Figure 5. As is evident from the diagram, there is little variation in average score for the stop consonants, but there are substantial variations among the fricative consonants. In particular, the dental fricatives ($\theta$ and $\delta$), showed very low scores, whereas the alveolar fricatives (s and z) showed by far the highest scores. Although voicing was not a significant main effect, there were significant interactions with voicing. The $C \times V$ interaction, shown in the bottom left section of Figure 5, was fairly complex. The data show that for two cognate pairs, $k/g$ and $\theta/\delta$, the voiceless component had a significantly higher test score, whereas for the $\theta/v$ cognate pair, the voiced
Table 1 Analysis of Variance: Consonant Recognition Scores

<table>
<thead>
<tr>
<th>Effect</th>
<th>Mean Square</th>
<th>Degrees of Freedom</th>
<th>F Ratio</th>
<th>Significance Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cognate pair (C)</td>
<td>13.37</td>
<td>5,50</td>
<td>17.4</td>
<td>p &lt; .001</td>
</tr>
<tr>
<td>Voicing (V)</td>
<td>0.41</td>
<td>1,10</td>
<td>2.3</td>
<td>Not significant</td>
</tr>
<tr>
<td>Enhancement (E)</td>
<td>8.17</td>
<td>1,10</td>
<td>33.5</td>
<td>p &lt; .001</td>
</tr>
<tr>
<td>Age group (A)</td>
<td>5.23</td>
<td>1,10</td>
<td>1.7</td>
<td>Not significant</td>
</tr>
<tr>
<td>C x V</td>
<td>0.98</td>
<td>5,50</td>
<td>4.4</td>
<td>p &lt; .005</td>
</tr>
<tr>
<td>C x E</td>
<td>0.36</td>
<td>5,50</td>
<td>3.6</td>
<td>p &lt; .01</td>
</tr>
<tr>
<td>E x A</td>
<td>1.01</td>
<td>1,10</td>
<td>4.2</td>
<td>p &lt; .1</td>
</tr>
<tr>
<td>V x E x A</td>
<td>0.53</td>
<td>1,10</td>
<td>6.0</td>
<td>p &lt; .05</td>
</tr>
<tr>
<td>Subjects</td>
<td>3.13</td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subject interactions</td>
<td>0.07 to 0.77</td>
<td>10 or 50</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The table summarizes the results of a five-factor, repeated-measures analysis of variance repeated over subjects. Two groups of six subjects participated so that the degrees of freedom (df) for the subject effect were reduced to 10 (1 df for each group). Similarly, the df for the subject interactions, used as the denominators in the various F-tests, are reduced to either 50, for interactions involving C, or 10, for interactions not involving C. All sums of squares are in (arcsine units)? since the arcsine transformation was used prior to statistical testing.

component scored significantly higher. The remaining cognate pairs showed no significant differences between the voiced and voiceless components.

Of particular interest in this study was the magnitude of the enhancement effects and the interactions of this effect with the other main effects. As shown in the top right section of Figure 5, the average test score increased by 16.5 percentage points, from 33.3 percent correct for P0 to 49.8 percent correct for Pmax. The magnitude of this improvement was found to differ significantly between age groups (i.e., there was a significant E x A interaction). As shown in the middle section of Figure 5, the enhancement effect was smaller for the younger age group, the average increase in test score being 10.9 percentage points (from 42.7% to 53.6% correct), whereas the improvement for the older age group was correspondingly larger. The increase in this case was 21.5 percentage points, from 24.5 percent to 46.0 percent correct. (Note that the unweighted average of the two percentage point improvements is \( \frac{1}{2} (10.9 + 21.5) = 16.2 \), which is slightly less than the weighted average of 16.5 percentage points. The use of the arcsine transformation in analyzing the data ensures that all averages are weighted appropriately for data in the form of percentages. The weighting used is inversely proportional to the error variance of each estimated percentage.)

There was also a significant interaction between voicing, enhancement, and age group. This three-way interaction (V x E x A) is shown in the bottom right section of Figure 5. The data show that the enhancement effect was largest for the voiceless consonants in the older children. The test score more than doubled in going from 24.4 percent correct for Po to 51.1 percent correct for Pmax, the increase in score being 26.7 percentage points. The next largest improvement of 16.3 percentage points was shown by the voiced consonants in the older children. The younger children showed a slightly smaller improvement of 14.2 percentage points for the voiced consonants and a relatively small improvement of 7.7 percentage points for the voiceless consonants.

Analysis of CEmax

A four-factor, repeated-measures analysis of variance was performed on the CEmax data. The factors were cognate pair (C), voicing (V), and age group (A) repeated over subjects (S). The results of the analysis are summarized in Table 2. The CEmax data show a much more complex pattern than the percent correct data in that all three main effects (C, V, and A) and all interactions except one (C x A) were statistically significant.

The top third of Figure 6 provides a summary of the main effects. The effect of cognate pair (C) on CEmax is relatively small for the stop consonants but quite substantial for the fricative consonants, from a low of 4.4 dB for 0/8/ to a high of 11.4 dB for s/z. The voiceless consonants, on average, required significantly larger values of CEmax (8.6 dB as compared to 5.8 dB for the voiced consonants). The differences in CEmax between age groups were even more pronounced, from an average of 3.9 dB for the younger age group to an average of 10.5 dB for the older age group.

The significant interactions shown by the CEmax data are summarized in the lower two-
Figure 5 The top third of the diagram shows percent correct scores for the significant main effects of cognate pair (C) and enhancement (E). Interactions C × E and E × A are shown in the middle third of the diagram and interactions C × V and V × E × A are shown in the bottom third of the diagram. Unenhanced (Po) conditions are represented by solid bars and the enhanced (Pmax) condition by open bars. The open bars are superimposed on the solid bars so that the open section above the solid area represents the gain in percent correct score.

thirsts of Figure 6. Although the interactions are fairly complex, several salient features of these interactions stand out. The older children not only showed much larger values of CEmax, on average, but the pattern of interactions was also more consistent. In particular, CEmax was
Table 2  Analysis of Variance: CEmax

<table>
<thead>
<tr>
<th>Effect</th>
<th>Mean Square</th>
<th>Degrees of Freedom</th>
<th>F Ratio</th>
<th>Significance Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cognate pair (C)</td>
<td>133.9</td>
<td>5,50</td>
<td>2.78</td>
<td>p &lt; .05</td>
</tr>
<tr>
<td>Voicing (V)</td>
<td>294.1</td>
<td>1,10</td>
<td>15.23</td>
<td>p &lt; .05</td>
</tr>
<tr>
<td>Age group (A)</td>
<td>1564.2</td>
<td>1,10</td>
<td>15.22</td>
<td>p &lt; .005</td>
</tr>
<tr>
<td>C x V</td>
<td>78.6</td>
<td>5,50</td>
<td>2.74</td>
<td>p &lt; .05</td>
</tr>
<tr>
<td>C x A</td>
<td>40.3</td>
<td>5,50</td>
<td>0.84</td>
<td>Not significant</td>
</tr>
<tr>
<td>V x A</td>
<td>346.0</td>
<td>1,10</td>
<td>17.91</td>
<td>p &lt; .005</td>
</tr>
<tr>
<td>C x V x A</td>
<td>56.6</td>
<td>5,50</td>
<td>1.98</td>
<td>p &lt; .1</td>
</tr>
<tr>
<td>Subjects</td>
<td>102.7</td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subject interactions</td>
<td>19.3 to 48.2</td>
<td>10 or 50</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The table summarizes the results of a four-factor, repeated-measures analysis of variance repeated over subjects. Two groups of six subjects participated. Since the experimental design was the same as that for the preceding analysis, the available degrees of freedom were the same as before. (See caption to Table 3.) Note, however, that in this case the unit of measurement was the dB and that the arcsine transformation was not used.

significantly higher for the voiceless component of each cognate pair, with the exception of /f/, which required less enhancement for Pmax than its voiced cognate /v/. The pattern for younger children was the same for the fricatives but not for the stops.

**DISCUSSION**

The most important finding of this investigation is that consonant recognition in young children can be improved significantly by increasing consonant levels. Although there were large differences between children in terms of their percent correct scores, there were nevertheless patterns of improvement that were similar across children and across age groups. It is also possible to compare these patterns of improvement with those of adults having similar sensorineural hearing losses. Kennedy et al (1998) used the same test stimuli with a wide range of adult listeners.

Table 3 shows the improvement in consonant recognition for stops and fricatives for each of the two age groups considered in this study, as well as for adults with similar hearing losses (adult data drawn from Kennedy et al, 1998). The improvements are specified in arcsine units in order to avoid problems relating to differences in the error variance of percentage scores. As is evident from the table, the improvements in consonant recognition did not differ significantly between the older children (8 to 9 years of age) and the adults, but the younger children (5 to 6 years of age) showed significantly smaller improvements. The children also differed from the adults in that the improvement in consonant recognition did not differ significantly between stops and fricatives for children, whereas the adults showed a significantly larger improvement for the fricatives. This result was contrary to expectation in that the processed (enhanced) stop consonants sound more distorted or less natural than the processed fricatives. It was expected that children, especially younger children, would have had greater difficulty than adults in recognizing consonants that sound more distorted as a result of the enhancement process. It may be that for the younger, hard-of-hearing child, any speech sound that has been processed so as to improve audibility may sound equally distorted.

Table 3 also shows average values of CEmax. The data show differences among the younger children, older children, and adults that are analogous to those obtained for consonant recognition. The average values of CEmax for the younger children were significantly less than those for the older children, and the values of CEmax for the older children did not differ significantly from adult values. The lower average values of CEmax for the younger children are largely a result of three younger children having relatively narrow dynamic ranges for CE. These children found enhancements of more than 6 or 9 dB uncomfortably loud. If the data for these children are omitted from the analysis, the relative improvement in consonant recognition for the younger age group approaches that for the older age group. One interpretation of this result is that young children develop the ability to deal with consonant enhancement (i.e., consonants with high gain) over time, and that some children develop this ability sooner than others.

There were other similarities and dissimilarities between the test scores for children and adults. Relative performance for the six fricative
Consonant Enhancement Effects on Speech Recognition

Smith and Levitt

COGNATE PAIR (C) VOICING (V) AGE GROUP (A)

\[ v = \text{voiceless} \]
\[ V = \text{voiced} \]
\[ Y = \text{younger} \]
\[ O = \text{older} \]

\[ C \times V \text{ INTERACTION} \]
\[ V \times A \text{ INTERACTION} \]

\[ C \times V \times A \text{ INTERACTION (YOUNGER AGE GROUP)} \]
\[ C \times V \times A \text{ INTERACTION (OLDER AGE GROUP)} \]

**Figure 6** Average values for each of the three significant main effects, cognate pair (C), voicing (V), and age group (A), are shown in the top third of the diagram. The C \times V and V \times A interactions are shown in the middle of the diagram and the three-way interaction, C \times V \times A, is shown in the bottom third of the diagram. The symbol \( v \) represents voiceless consonants and \( V \) represents voiced consonants. Similarly, \( Y \) represents younger age group and \( O \) represents the older age group.

Consonants was much the same for both children and adults. Consonant scores were ranked for each age group and compared to the ranking obtained by the adults. The Spearman rank correlation coefficient, rs, was then computed. As shown in Table 3, rs was close to 1.0 for the fricative consonants for both the unenhanced (Po) and enhanced (Pmax) conditions, indicating essentially the same ranking of these consonantal sounds for these age groups. In contrast, the ranking of scores for the stop consonants is quite different between children and adults. None of the Spearman rank correlation coefficients were significant for the stop consonants. Part of the reason may have been the smaller range of scores for the stop consonants, which
Table 3  Comparisons with Adult Data

<table>
<thead>
<tr>
<th></th>
<th>Children 5–6 Years of Age</th>
<th>Children 8–9 Years of Age</th>
<th>Adults 20–80 Years of Age</th>
</tr>
</thead>
<tbody>
<tr>
<td>Improvement in Consonant Recognition</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fricatives (arcsine units)</td>
<td>0.17</td>
<td>0.49</td>
<td>0.62</td>
</tr>
<tr>
<td>Stops (arcsine units)</td>
<td>0.26</td>
<td>0.42</td>
<td>0.38</td>
</tr>
<tr>
<td>Standard error of mean (arcsine units)</td>
<td>0.06</td>
<td>0.06</td>
<td>0.08</td>
</tr>
<tr>
<td>Average Value of CEmax</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fricatives (dB)</td>
<td>3.6</td>
<td>11.2</td>
<td>10.8</td>
</tr>
<tr>
<td>Stops (dB)</td>
<td>4.2</td>
<td>9.8</td>
<td>12.3</td>
</tr>
<tr>
<td>Standard error of mean (dB)</td>
<td>1.1</td>
<td>1.1</td>
<td>1.3</td>
</tr>
<tr>
<td>Spearman Rank Correlation Coefficient</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Correlations with Adult Data)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fricatives</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unenhanced</td>
<td>0.94**</td>
<td>0.89*</td>
<td>1.00</td>
</tr>
<tr>
<td>Enhanced</td>
<td>0.94**</td>
<td>0.83*</td>
<td>1.00</td>
</tr>
<tr>
<td>Stops</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unenhanced</td>
<td>0.03</td>
<td>0.26</td>
<td>1.00</td>
</tr>
<tr>
<td>Enhanced</td>
<td>0.60</td>
<td>-0.09</td>
<td>1.00</td>
</tr>
</tbody>
</table>

The table provides several comparisons with the data of Kennedy et al (1998) for the same test stimuli evaluated with adults having similar sensorineural hearing losses. Improvements in consonant recognition for stops and fricatives are shown as well as average values of CEmax for these consonant groups. Also shown are Spearman rank correlations between children and adult consonant recognition scores. Arcsine units are used in comparing consonant recognition scores in order to avoid problems relating to differences in error variance at different percentage values.

*Significant at .05 level; **significant at .01 level.

reduces the likelihood of similar rankings. Another possible reason is that stop consonants are more complex than fricative consonants and, as a consequence, there are greater individual and developmental differences in the perception of stop consonants.

The results of this investigation have several important implications for the design of hearing aids for children and several important caveats. The data indicate that, if appropriate gain can be provided to consonantal sounds in speech, large improvements in intelligibility should result, particularly for children over 8 years of age. As noted earlier, the average consonant recognition score for the older children increased from 24.5 to 46.0 percent. This improvement in score for nonsense syllables corresponds roughly to an improvement from 60 to 90 percent for everyday sentences in adults (Kryter, 1985). This is a substantial improvement and would be of great clinical value if it could be realized in a practical hearing aid. Even the smaller improvement obtained with the younger children represents a substantive gain in consonant recognition and, with appropriate auditory training, the gains in overall intelligibility could be equally impressive for younger children.

The problem of automatically adjusting gain to CEmax in running speech is particularly difficult, and the improvements obtained in this study should be viewed as an upper bound that could be achieved if techniques could be developed for making the appropriate gain adjustments automatically and with minimal processing distortions. Since the potential improvement in intelligibility is substantial, this represents a worthwhile avenue of investigation.

Another caveat to bear in mind is that the data were obtained with nonsense syllables as opposed to continuous speech. There are, of course, important differences between nonsense syllables and continuous speech, but as shown by French and Steinberg (1947), data obtained with nonsense syllables can be quite effective in predicting the intelligibility of continuous speech.

Despite these caveats, there are aspects of the data that are very promising with respect to implementing automatic adjustment of CEmax in a practical hearing aid. Sensory aids have already been developed that extract voiceless fricatives automatically in continuous speech (Guttman and Nelson, 1968). Further, CE functions for the voiceless fricatives were mostly flat or gradually rising; therefore, CEmax can be approximated quite well for these sounds in a practical instrument. CE functions with a sharp peak present a problem for automatic adjustment, but these CE functions were fairly infrequent, occurring only about 10 percent of
the time. Voiced fricatives are more difficult to
\text{detect automatically than voiceless fricatives, but
with appropriate algorithms they can be detected
with a relatively low error rate. Automatic detec-
tion of the stop consonants, unfortunately, is a
much more difficult problem.}

Weak sounds with high-frequency content
can also be made more audible by the simple
expedient of combining a high-frequency boost
with amplitude compression. The problem here
is that of selecting the right time constants for
continuous speech. The fricative consonants, as
before, are well suited for this type of process-
ning and signal-processing strategies of this type
could be implemented in a modern digital hear-
ing aid (Preves, 1994). It remains to be seen
which approach yields the best results. The data
provided by this study should serve as a useful
benchmark in evaluating the efficacy of the var-
ious practical approaches to consonant enhance-
ment that are currently being explored.

\textbf{Acknowledgment.} The research reported in this paper
was supported by NIH-NIDCD Grant No. 5P 50 DC00178.

\section*{REFERENCES}

Abromovitz A. (1971). Assessment of auditory percep-

Frequency resolution in children. \textit{J Speech Hear Res}
32:317–322.

Bennet CA, Franklin NL. (1954). \textit{Statistical Analysis in
Chemistry and the Chemical Industry}. Section 6.32. New
York: Wiley.

Dubno JR, Levitt H. (1981). Predicting consonant con-
fusions from acoustic analysis. \textit{J Acoust Soc Am}

French NR, Steinberg JC. (1947). Factors governing the
intelligibility of speech sounds. \textit{J Acoust Soc Am}
19:90–119.

consonant-vowel ratio modification on amplitude envelope
cues for consonant recognition. \textit{J Speech Hear Res}

on consonant perception by elderly hearing-impaired sub-

Guttmann N, Nelson JR. (1968). An instrument that cre-
ses some artificial speech spectra for the severely hard

Irwin RJ, Stillman JA, Shade A. (1986). The width of the

Consonant-vowel intensity ratios for maximizing conson-
ant recognition by hearing-impaired listeners. \textit{J Acoust
Soc Am} 103:1098–1114.


Levitt H, Resnick S. (1978). Speech reception by the
hearing-impaired: methods of testing and the develop-

enhancement techniques to improve intelligibility for


Picheny MA, Durlach NI, Braida LD. (1986). Speaking
clearly for the hard of hearing II: acoustic characteris-
tics of clear and conversational speech. \textit{J Speech Hear
Res} 29:434–446.

Preves D. (1994). Current and future applications of
digital hearing aid technology. In Sandlin RE, ed. \textit{Under-
standing Digitally Programmable Hearing Aids}. Needham
Heights, MA: Allyn & Bacon.

Some rehabilitative considerations for future speech-


Psychoacoustics}. Washington, DC: American Psychological
Association.