Benefits of amplification for speech recognition in background noise

Christopher W. Turner and Belinda A. Henry
Department of Speech Pathology and Audiology, University of Iowa, Iowa City, Iowa 52242

(Received 29 January 2002; accepted for publication 12 July 2002)

The purpose of the present study was to examine the benefits of providing audible speech to listeners with sensorineural hearing loss when the speech is presented in a background noise. Previous studies have shown that when listeners have a severe hearing loss in the higher frequencies, providing audible speech (in a quiet background) to these higher frequencies usually results in no improvement in speech recognition. In the present experiments, speech was presented in a background of multitalker babble to listeners with various severities of hearing loss. The signal was low-pass filtered at numerous cutoff frequencies and speech recognition was measured as additional high-frequency speech information was provided to the hearing-impaired listeners. It was found in all cases, regardless of hearing loss or frequency range, that providing audible speech resulted in an increase in recognition score. The change in recognition as the cutoff frequency was increased, along with the amount of audible speech information in each condition (articulation index), was used to calculate the ‘‘efficiency’’ of providing audible speech. Efficiencies were positive for all degrees of hearing loss. However, the gains in recognition were small, and the maximum score obtained by an listener was low, due to the noise background. An analysis of error patterns showed that due to the limited speech audibility in a noise background, even severely impaired listeners used additional speech audibility in the high frequencies to improve their perception of the ‘‘easier’’ features of speech including voicing. [DOI: 10.1121/1.1506158]

PACS numbers: 43.71.Ky, 43.71.Gv, 43.66.Sr [KRK]

I. INTRODUCTION

There have been a considerable number of recent studies suggesting that the benefits of providing audible speech to listeners with sensorineural hearing loss (as is done by a hearing aid), has limitations. For example, Ching *et al.* (1998) and Hogan and Turner (1998) showed that providing bands of high-frequency speech at audible levels for severe hearing losses often resulted in no increase or even a decrease in speech recognition for some patients. This trend was evident for speech information above approximately 2500–3000 Hz. Turner and Cummings (1999), Skinner (1980), and Rankovic (1991) provided similar evidence that maximizing the amount of audible speech was not always the most beneficial strategy for patients with sensorineural hearing loss. On the other hand, Ching *et al.* (1998) and Turner and Brus (2001) demonstrated that for lower-frequency regions of the speech range, amplifying speech to audible levels consistently provided benefit to all patients.

This linking of a limited benefit for amplification to the degree of hearing loss implies that the degree and type of cochlear damage is an important factor in determining whether audible speech will be beneficial or not. A number of authors (e.g., Van Tasell, 1993; Turner, 1999; Turner and Cummings, 1999; Turner and Brus, 2001; Vickers *et al.*, 2001) hypothesized that these severe hearing losses indicate damage to the inner hair cells. Vickers *et al.* (2001) have also shown limited benefits of amplification and suspected that this occurs when speech is presented to ‘‘dead regions’’ of the cochlea. Such damage could interfere with the perception of speech, particularly with the place of articulation feature of speech, which is carried to a large extent by the higher-frequency regions of speech.

All of the above-mentioned studies looked at the recognition of speech in quiet (i.e., no background noise). In the present study we extended this line of research to the situation where speech is presented in a background noise that has a spectrum similar to the speech. The present study determined the ability of hearing-impaired listeners, with a range of hearing loss degrees and configurations, to extract speech information from the audible portion of the speech signal when listening in background noise. At the outset of this study, we quite expected that our results would lead to conclusions similar to the previous research. Since it is well known that listeners with sensorineural hearing loss often have difficulty understanding speech in noisy backgrounds (even when they may do quite well in quiet), our expectations were that the benefits of amplification would again be zero or even negative in some situations (severe loss and high frequencies), perhaps even to a greater extent than in the previous research which measured speech recognition in quiet. However, our original predictions were not realized. In contrast, we found that amplification of speech in a background noise always provided some benefit for listeners with hearing loss, regardless of the degree of hearing loss and/or the frequency region of speech. Although these gains in speech recognition were for the most part small, this contrary finding does provide additional insights into the speech recognition abilities of listeners with sensorineural hearing impairment. It also can serve as a caution in accepting the con-

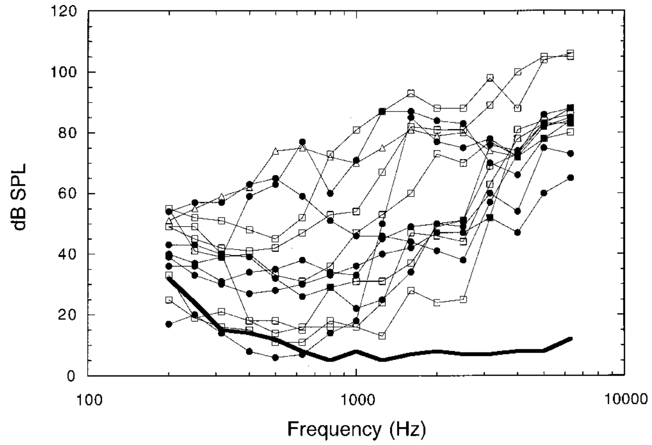

FIG. 1. The pure-tone sensitivity thresholds for the normal-hearing and hearing-impaired subjects in this study. The heavy solid line represents the average thresholds of the normal-hearing subjects. The lighter lines with individual symbols show the thresholds of the hearing-impaired listeners.

clusions of the previous research in too sweeping of a manner.

II. METHODS

A. Subjects

Five normal-hearing listeners and 13 listeners with sensorineural hearing loss were recruited for this study. The normal-hearing subjects had pure-tone sensitivity thresholds better than 20 dB HL (ANSI, 1996) at all octave test frequencies from 250 to 8000 Hz. The 13 listeners with sensorineural hearing loss were specifically recruited to yield a sampling of various degrees of sensorineural hearing loss across the frequency range. The pure-tone thresholds for these subjects are displayed in Fig. 1. The bold line shows the average thresholds for the group of normal-hearing subjects. Each of the lines with symbols represents the thresholds for an individual hearing-impaired subject. All subjects were native speakers of American English. All of the hearing-impaired listeners had bilateral hearing losses and the better ear of each was used for all testing.

B. Stimuli

Pure-tone thresholds were measured at the center frequencies of one-third-octave bands from 200 to 8000 Hz for calculations of speech audibility using the articulation index (AI). Test tones were 500 ms in duration with 25-ms rise–fall times. All testing (thresholds and speech recognition) was done in a sound booth using Sennheiser HD 25-SP supra-aural headphones. All sound levels reported in these experiments are referenced to the levels developed by these headphones in the NBS-9A coupler.

The speech recognition testing in this experiment used the same materials as in our previous studies (Hogan and Turner, 1998; Turner and Brus, 2001). The 12 lists of the Nonsense Syllable Test (NST, UCLA version) were used to measure consonant recognition. These consonants are well-suited to subsequent error pattern analysis. Each of the 12 lists consisted of 21 or 22 consonants presented with a fixed talker (male or female), consonant position (initial or final), and vowel context. One hundred randomly chosen stimuli were presented from each list; thus, a data point for an individual subject in a particular listening condition was based upon 1200 trials. All speech tokens were stored digitally and presented under computer control (Macintosh G4) through a 16-bit digital-to-analog converter (Audiomedia III, Digidesign, Inc.) at a sampling rate of 44.1 kHz with a built-in antialiasing filter set to 20 kHz. For the hearing-impaired subjects, the speech materials were presented through an analog high-pass emphasis spectrum shaper (Altec-Lansing 1753). Several versions of high-pass shaping were employed, as well as a range of presentation levels, chosen to accommodate the variation in hearing-loss configurations across subjects. The high-pass shaping provided from 15–30 dB of relative gain for frequencies above 1000 Hz.

C. Procedures

Pure-tone thresholds were measured using a computer-controlled adaptive procedure that varied tone levels via a Tucker-Davis programmable attenuator (model PATT). A one-up, two-down, four-alternative forced-choice procedure was employed using 2-dB steps and the threshold was taken as the average of the final 8 of 13 reversals of the procedure.

For speech testing, the consonant phonemes were displayed as labeled buttons on a touchscreen (MicroTouch). The subjects were instructed to respond following the presentation of each token by touching the corresponding button on the touchscreen. All subjects first participated in several practice sessions in which they were given feedback and learned to associate the various tokens with the correct consonant response button. For these conditions, speech was presented in a wideband condition (low-pass filtered at 9000 Hz) with no background noise present. Normal-hearing listeners heard the speech at 70 dB SPL without high-pass shaping. The subjects with hearing loss listened to the speech through one of the spectrum shapers, which was chosen on an individual basis, in an attempt to maximize the frequency range for which speech could be made audible. An appropriate presentation level of the speech was also determined at this time by asking the subject to choose the level that ''provided the most information about the speech sounds yet was not uncomfortably loud.'' The chosen spectral level of speech for each subject was then used for a series of low-pass filtered conditions, described below. For several of the more severely hearing-impaired listeners, speech recognition was also obtained at one additional higher speech presentation level in a further attempt to provide maximum speech audibility at the highest frequencies.

Each subject's recognition score for the NST materials was then measured for the broadband condition without background noise. No trial-by-trial feedback was given to the subject for this testing, or for any subsequent testing. The recognition scores in quiet for all the normal-hearing subjects was at least 96%; for the hearing-impaired subjects the mean recognition scores in quiet were 65% (range 51 to 77%).

The background noise was a multitalker babble consisting of both male and female voices played continuously from a compact disk recording throughout each testing session. It

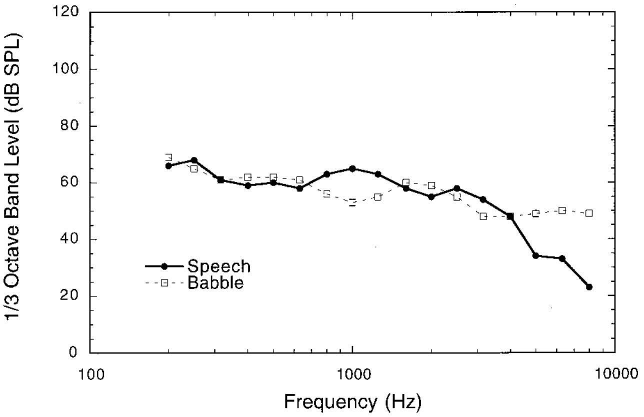

FIG. 2. The long-term average spectra of the NST speech lists and the multitalker babble. For this figure the signals were presented at equal overall sound-pressure levels and were not presented through the spectrum shaper.

was attenuated and mixed with the speech channel prior to any spectrum shaping. Figure 2 displays the long-term spectra of both the concatenated NST stimuli (with no silent spots between tokens) and the babble, as measured at the output of the headphones. For each subject, an appropriate level of background noise was then chosen during some additional pilot testing and was then used for the remainder of the speech recognition testing. Our goal in setting the background noise level was to reduce the subjects' scores in the noise to be approximately two-thirds of their score in quiet, thus providing a consistent reduction of score across subjects. For example, if a subject's score in quiet was 60%, the level of the background noise was chosen to yield a recognition score of approximately 40%. For the normal-hearing listeners, a signal-to-babble ratio of +4 dB provided the appropriate decrease in recognition, and was used for all normal-hearing subjects. The average signal-to-babble ratio for the hearing-impaired subjects was +9 dB (range +4 to +14 dB).

Speech plus background noise was presented to each subject under seven low-pass filtering conditions. The low-pass filter cutoffs were 350, 560, 900, 1400, 2250, 3500, and 5600 Hz, as well as the broadband (9000 Hz) condition. A Kemo (VBF8.04) filter with slopes of 30 dB/octave was used. The order of filter conditions was randomly ordered for each subject.

III. RESULTS

A. Data analysis

The data were analyzed using a method identical to that used in our previous studies (Hogan and Turner, 1998; Turner and Brus, 2001), and the reader is referred to those studies for a more detailed description of the procedures. For each filter cutoff condition, a final recognition score based upon the average of the 12 lists was obtained. A value of articulation index (AI) or speech intelligibility index (SII) was calculated for each listening condition for each subject (ANSI, 1969, R1997) using the subjects' pure-tone thresholds, and the presentation levels of the filtered speech and background noise in that condition. The one-third-octave band AI method using the frequency-importance function for

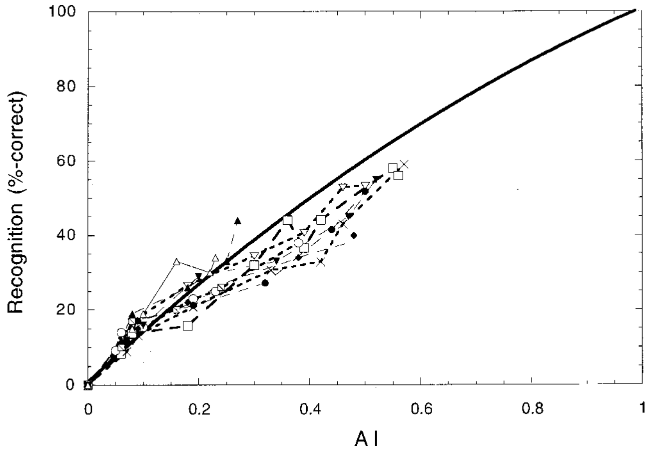

FIG. 3. Recognition in percent correct as a function of the amount of audible speech information (AI) for all subjects. The solid line represents a curve fitted to the normal-hearing subjects's data. The individual lighter lines with symbols represent individual hearing-impaired subjects' data.

these NST materials was employed. The calculated AI is a value between 0.0 and 1.0 representing the proportion of speech formation available to the listener. Data analysis using the earlier (1969) version is presented in this report; however, we also performed the analysis using the newer version (R1997) and the conclusions are essentially the same. The data were plotted as speech recognition in percent-correct) as a function of the degree of audible speech information (AI). The raw data (percent correct as a function of AI) for all hearing-impaired listeners are displayed in Fig. 3, along with the average articulation function for the normal-hearing listeners. The data for all the normal-hearing listeners were pooled and fit with a second-order polynomial to serve as a reference. The data of the normal-hearing subjects in this study were essentially identical to the normal-hearing subjects of Turner and Brus (2001). In general, all the functions for hearing-impaired subjects show increases in recognition as AI increases.

A second-order polynomial function was also fit to each individual hearing-impaired subject's data. See Fig. 4 for two

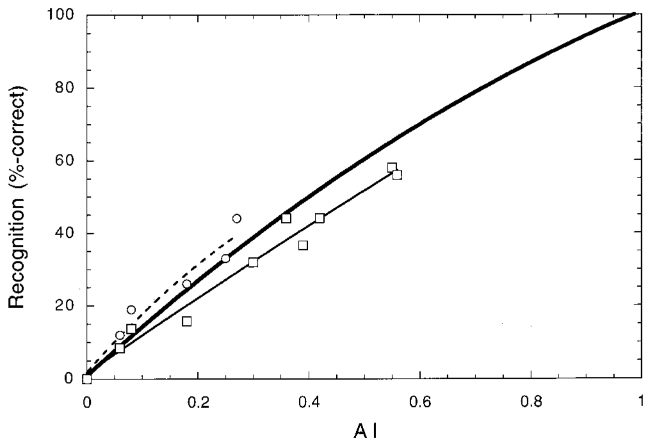

FIG. 4. The articulation functions for the normal-hearing subjects and also for two of the hearing-impaired subjects. The heavy solid line is the fitted curve to the data of the normal-hearing subjects. The two symbol types display the data from the two hearing-impaired subjects. The lighter lines are the fitted curves for those two hearing-impaired subjects.

examples. The dark solid line of Fig. 4 represents the fitted articulation function of the normal-hearing subjects in this study. As the low-pass filter cutoff was increased to 3500 Hz and above, not all hearing-impaired subjects obtained an increase in their calculated audible speech information (AI), particularly if their pure-tone sensitivity thresholds were highly elevated for those higher frequencies (for example, the subject represented by the open circles in Fig. 4). Those data points were not included in this figure or the data analysis. And, the curve fit was based upon fewer than seven data points. For other subjects, as mentioned previously, consonant recognition was measured at an additional higher presentation level, yielding additional data points (for example, the open squares of Fig. 4).

As in our previous studies, the question asked was ''For a given increment of audible speech information presented to a hearing-impaired subject, how will the change in that subject's speech recognition compare to that of a normal-hearing subject receiving the same increment of audible speech information?'' In order to quantify this, we again used the measure of ''efficiency,'' which is simply the ratio of the hearing-impaired listener's recognition improvement to the normal-hearing listener's improvement, measured at the same AI value on each subject's articulation function. All measures of efficiency were calculated from the subjects' fitted curves, as in our previous studies. An efficiency of 1.0 means that for this increment of speech audibility, the listener used the newly audible speech just as well as a normal-hearing listener would. An efficiency of 0.0 means that the listener received no benefit from the audible speech.

The subjects' pure-tone thresholds were known for each one-third-octave band test frequency, allowing one to relate the efficiency for each increment of audible speech to the degree of hearing loss present at the frequencies of the increment. These sensitivity thresholds could then be compared to the values of a large group of normal-hearing listeners who were measured with the same Sennheiser HD-25SP headphones (Hogan and Turner, 1998), to yield the degree of hearing loss. Our increments of speech audibility, obtained by increasing the low-pass cutoff frequency of the filtered speech, were 2 one-third-octave bands wide. For our measure of the degree of hearing loss, the sensitivity threshold was taken as the average of the two bands.

In Fig. 5, the efficiencies of all hearing-impaired subjects for all frequencies are displayed as function of the degree of hearing loss. In every case, for hearing losses up to 90 dB HL, the calculated efficiency was positive, indicating that providing additional audible speech to patients with all degrees of hearing loss provides benefits in speech recognition for speech presented in a background noise.

Of particular interest here are the efficiencies for the higher-frequency regions of speech. Several previous studies, as mentioned above, found negative or zero benefits of audible speech (under quiet listening conditions) for higher-frequency regions of speech when the hearing loss was severe. In Fig. 6, the calculated efficiencies for the speech bands of 2250–3500 Hz and also for 3500–5600 Hz are plotted separately, thus indicating the benefits of audible speech for frequencies of 2250 Hz and above. No data were available for the speech band of 5600–9000 Hz, as audible speech could not be provided to any of the hearing-impaired listeners for that frequency range without exceeding uncomfortable loudnesses for the subjects. As expected from the previous figure, all efficiencies are positive. When linear regression lines are fit to the data points of these higher frequencies, the extrapolated intersections with a value of zero efficiency are over 110 dB HL in each case, further suggesting that providing audible speech to any degree of hearing loss does provide benefits for speech recognition in a background noise.

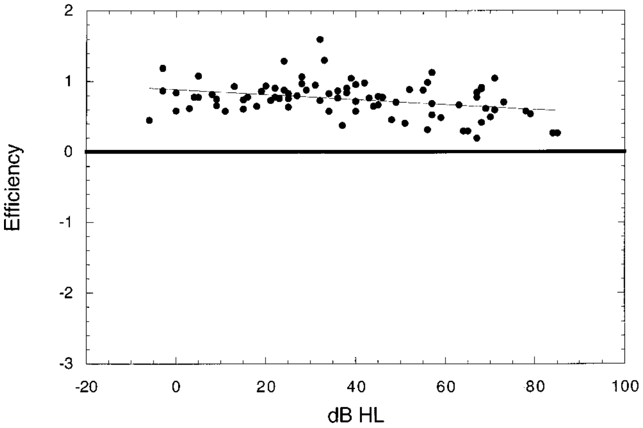

FIG. 5. The efficiency of audible speech for recognition plotted as function of the degree of hearing loss. All subjects and all conditions are shown in this figure. The degree of hearing loss is expressed as the difference between the pure-tone thresholds of normal-hearing subjects and the pure-tone threshold of the hearing-impaired subject.

IV. DISCUSSION

In the present study, it was shown that for listening to speech in a substantial background noise, all hearing-impaired listeners obtained benefit from amplified audible speech regardless of the degree of hearing loss. This finding held true even for the higher-frequency regions of the speech spectrum. This appears to be in contrast to the conclusions of the previous research Ching *et al.* (1998) and Hogan and Turner (1998), as well as others. The primary difference between the present study and these previous ones is that

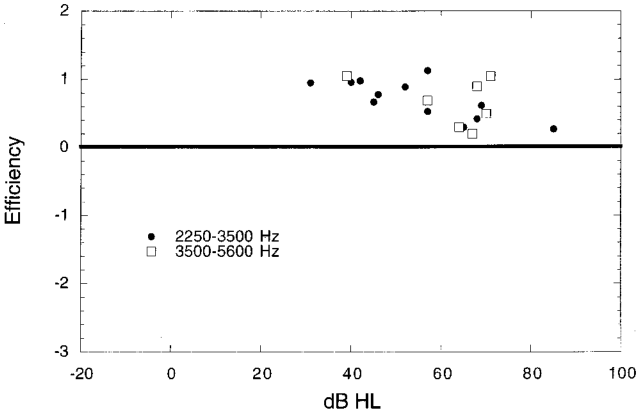

FIG. 6. The efficiency of audible speech for recognition plotted as a function of the degree of hearing loss. In this figure, only the data for low-pass filter cutoff frequencies of 3500 and 5600 Hz are shown.

speech recognition was measured in a background noise, and the previous work measured speech in a quiet background. Thus, our original hypotheses were not confirmed; the benefits of amplified speech in a background noise were not zero or negative for high frequencies and severe hearing loss. We had also hypothesized that these deficits would be more severe in noise backgrounds than in the previously measured effects in quiet; this certainly was not the case.

An examination of the raw data of the present study, as well as that of previous studies, provides a possible explanation for the present results. In the study of Hogan and Turner (1998), the hearing-impaired subjects obtained speech recognition scores ranging from 40% to 90% in the broadest bandwidth conditions, and it was typical for the presented speech (with no background noise present) to yield maximum AI values ranging from approximately 0.6 to 0.8 in these conditions. In other words, in the previous research, the majority of the speech signal was audible to the hearing-impaired listeners. In the present study, due to the spectrally similar background noise, the hearing-impaired subjects' speech recognition scores ranged from 24% to 55% in the broadest bandwidth conditions, and the speech signal was much less audible, with maximum AI values ranging from 0.23 to 0.62 (see Fig. 3). Thus, the earlier studies were measuring the ability of hearing-impaired listeners using the higher frequencies of speech to increase their speech recognition scores above values that were already rather high, presumably looking at their ability to recognize the remaining "most difficult" features and phonemes of speech. In the present study the additional high-frequency audible speech added was used by the hearing-impaired subjects to increase rather low recognition scores in every case. The noise background allowed these subjects only a small fraction of the possible audible speech in the lower-frequency conditions. As higher-frequency regions were made audible to these subjects, they presumably used the additional audible speech to increase recognition of "easier" features and phonemes of speech.

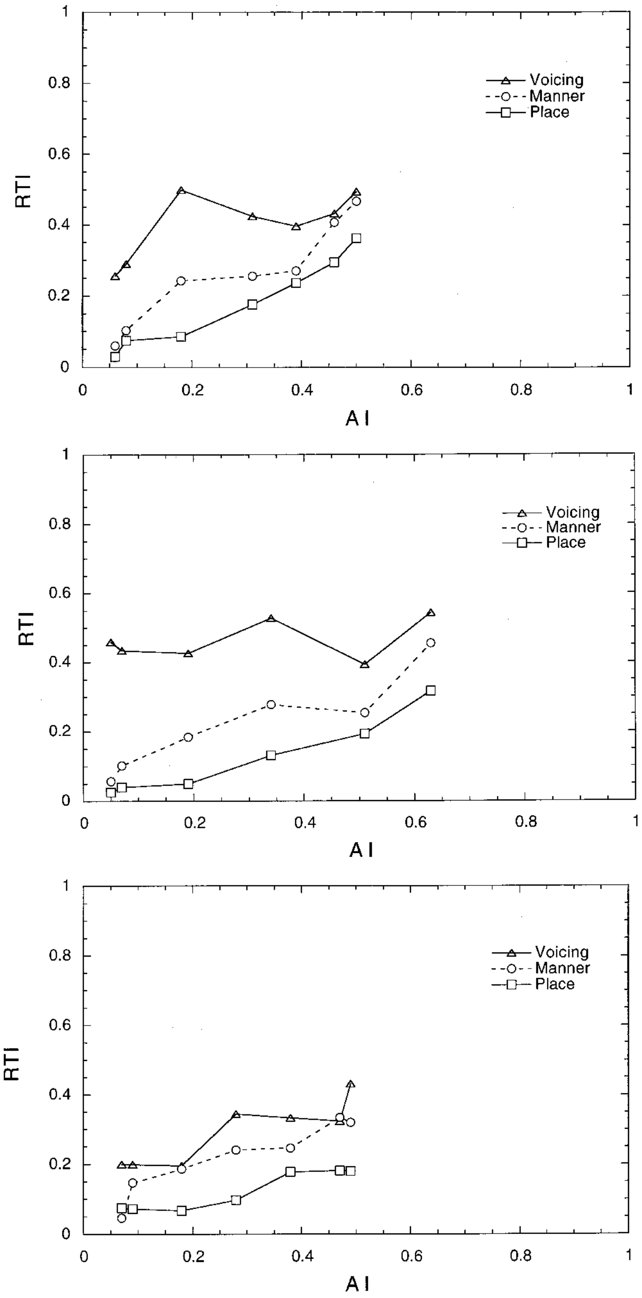

FIG. 7. The relative information transmitted (RTI) of the individual features of speech (voicing, manner, and place) is plotted as function for AI for three hearing-impaired subjects. Each panel corresponds to an individual subject.

One way to look at which types of speech features were being recognized by the subjects is to look at their perception of the commonly used distinctive features of speech such as voicing, manner, and place of articulation. For each hearing-impaired subject, the raw response matrices for speech recognition for each condition were analyzed using the FIX analysis program (Department of Phonetics and Linguistics, University College of London), which provides the relative information transmitted (RTI) measure of the transmission of these distinctive features of speech. This program is based upon the "sequential information analysis" described by Wang and Bilger (1973). The information transmitted was calculated iteratively, holding the order of analysis fixed as voicing, then manner, followed by place. In the three panels of Fig. 7, the RTI for voicing, manner, and place are displayed as a function of audible speech information (AI) for three of the hearing-impaired subjects. The trends seen in the data of these three subjects are similar to that in the other subjects' data. As one moves from the left to the right on each graph, additional high frequencies of speech have been made audible to the listener. For each subject, the values for voicing and manner are generally increasing even as speech is added up to 5600 Hz (the rightmost data point). And, in all three of the subjects, perception of the place of articulation feature is also increasing. In every instance, the RTI for any of the three features of speech reaches a maximum of 0.6 or less. In other words, even at the widest bandwidth condition the hearing-impaired subjects had plenty of speech cues remaining in which to show improvement. The fact that voicing cues are available across a wide bandwidth of speech has been shown by Grant and Walden (1996). Under quiet conditions, most listeners can get these same voicing cues from low-frequency regions of speech. In background noise, all frequency regions are used. The features of manner and place, which are usually associated with higher-frequency regions, also increased as listening bandwidth was increased. The results of the present study suggest that there are some easier cues for manner and even place that can be perceived by listeners with even severe hearing loss in the higher-frequency regions. This pattern of results is different from the feature analyses observed in previous research (speech in quiet) such as Vickers *et al.* (2001), where RTI for voicing and manner were often at 0.8 or above for wideband speech recognition. In these cases, only the more difficult cues of speech remained for the hearing-impaired listener to receive from increasing audibility at high frequencies, and apparently severe hearing loss can degrade the transmission of these remaining difficult speech cues.

It should be noted that the recognition gains these hearing-impaired subjects showed in response to the additional high-frequency bands of speech were usually not large (often on the order of 5% or less). These gains in speech recognition were also usually less than would be expected from normal-hearing listeners receiving equivalent increments in speech audibility, which further supports the idea that a hearing aid does not fully restore speech understanding for these patients. The small gains were also consistent with the small increments of AI that were added in these conditions (due to the presence of background noise and hearing loss). Although these subjects did not complain of uncomfortable loudness in these experimental conditions, listening to amplified speech and noise at these levels in everyday life may not be as acceptable. Thus, the clinical utility of providing large amounts of high-frequency gain should be viewed somewhat cautiously, despite the numerical conclusions of the present study.

In summary, the present study revealed that when the cues of speech are severely limited by background noise, providing audible speech via amplification showed positive benefit in all cases. Adding high-frequency audible speech in the presence of a relatively intense background noise was beneficial in all cases, regardless of the degree of hearing loss. In this case, when speech audibility is so limited, even the easier cues of speech remain elusive to the listener, and any small increase in speech audibility is put to good use.

ACKNOWLEDGMENTS

This research was funded in part by a grant from the National Institute of Deafness and other Communicative Disorders (1R01 DC00377). Many thanks to Hope Jones, Melanie Buhr, Courtney Burke, Ann Torkelson, Kristine Melis, and Arik Wald for their assistance on this project.

ANSI (**1969**). ANSI S3.5-1969, ‘‘American National Standard: Methods for the Calculation of the Articulation Index’’ (American National Standards Institute, New York).

ANSI (**1996**). ANSI S3.66-1996, ‘‘American National Standard: Specifications for Audiometers’’ (American National Standards Institute, New York).

ANSI (**R1997**), ANSI S3.6-1997, ‘‘American National Standard: Methods for the Calculation of the Articulation Index’’ (American National Standards Institute, New York).

Ching, T., 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.

Grant, K. W., and Walden, B. E. (**1996**). ‘‘Evaluating the articulation index for auditory-visual consonant recognition,’’ J. Acoust. Soc. Am. **100**, 2415–2424.

Hogan, C., and Turner, C. W. (**1998**). ‘‘High-frequency amplification: Benefits for hearing-impaired listeners,’’ J. Acoust. Soc. Am. **104**, 432–441.

Rankovic, C. M. (**1991**). ‘‘An application of the Articulation Index to hearing-aid fitting,’’ J. Speech Hear. Res. **27**, 391–402.

Skinner, M. W. (**1980**). ‘‘Speech intelligibility in noise-induced hearing loss: Effects of high-frequency compensation,’’ J. Acoust. Soc. Am. **67**, 306–317.

Turner, C. W. (**1999**). ‘‘The limits of high-frequency amplification,’’ Hear. J. **52**, 10–14.

Turner, C. W., and Brus, S. (**2001**). ‘‘Providing low- and mid-frequency speech information to listeners with sensorineural hearing loss,’’ J. Acoust. Soc. Am. **109**, 2999–3006.

Turner, C. W., and Cummings, K. J. (**1999**). ‘‘Speech audibility for listeners with high-frequency hearing loss,’’ Am. J. Audiol. **8**, 47–56.

Van Tasell, D. J. (**1993**). ‘‘Hearing loss, speech and hearing aids,’’ J. Speech Hear. Res. **36**, 228–244.

Vickers, D. A., Moore, B. C. J., and Baer, T. (**2001**). ‘‘Effects of low-pass filtering on the intelligibility of speech in quiet for people with and without dead regions at high frequencies,’’ J. Acoust. Soc. Am. **110**, 1164–1175.

Wang, M. D., and Bilger, R. C. (**1973**). ‘‘Consonant recognition in noise, a study of perceptual confusions,’’ J. Acoust. Soc. Am. **54**, 1248–1266.