Humes, Espinoza-Varas, and Watson (1988) recently reviewed much of the literature on the perception of sound by hearing-impaired listeners. They demonstrated that a model, subsequently referred to as a power-law model with compressed internal noise (Humes & Jesteadt, 1989), could account for most of the data from masking-based measures of frequency resolution in the hearing impaired. It was unclear, however, whether the performance of impaired listeners on masking-based measures of temporal resolution could also be accounted for by this model. The model maintains that the performance of listeners with sensorineural hearing loss on masking-based psychoacoustic tasks can be modelled by the introduction of an external masking noise into the ear of a normal listener. The latter masker is shaped to produce masked thresholds that approximate the audometric configuration of the impaired listeners. The modified power-law model is then used to predict the combined effect of two maskers: one masker used in the masking-based psychoacoustic measure and the other masker used to simulate the hearing loss of the impaired listener.

The evaluation of the model for masking-based measures of temporal processing was hampered by the scarcity of data to be analyzed. Consequently, the present experiment was conducted in which a masking paradigm was used to measure the preservation of modulation depth within the auditory system. Masking of pure tones by modulated noise has been measured previously in normal and impaired listeners by Zwicker and Schorn (1982) using the same stimuli and paradigm for both subject groups. Their results, obtained only for square-wave modulation of 15.6 Hz, suggest that impaired listeners have poorer temporal resolution than normal listeners with or without simulated hearing loss. This study, in part, seeks to replicate that finding.

Another interest in the present study was to provide a detailed description of the preservation of modulation within the auditory system for a range of modulation frequencies spanning those most prominent in running speech. Temporal-processing deficits observed previously in hearing-impaired listeners have been suggested frequently as significant contributors to the speech-understanding difficulties experienced by these listeners (Dreschler & Plomp, 1985; Irwin & McAuley, 1987; Tyler, Summerfield, Wood, & Fernandes, 1982). It is implied that such deficits are a second and independent manifestation of cochlear pathology, separate from the primary deficits of threshold elevation and accompanying loudness recruitment. The temporal-processing measures, however, frequently assess the limits of resolution in the time domain, as in the measurement of the minimum detectable temporal gap or the cut-off frequency of low-pass psychophysical modulation transfer functions (PMTFs). Resolution limits measured with the various PMTF procedures, for example, are often expressed as low-pass filter cut-off frequencies of 60-100 Hz, corresponding to time constants of 1–3 ms (Rodenburg, 1977; Viemeister, 1973, 1979). Modulation of the envelope of running speech, however, typically spans a much lower range of modulation frequencies. The envelope spectrum or modulation spectrum for running speech, for example, peaks at modulation frequencies of 3–4 Hz and shows little modulation beyond 20 Hz (Hori, House, & Hughes, 1971; Houtgast & Steeneken, 1985). Relatively little is known about the preservation of modulation in the auditory system of normal and impaired listeners for the range of modulation frequencies most prevalent in running speech.

In the present study, maskers having speech-like long-term average temporal and spectral characteristics were
used to explore the preservation of modulation within the auditory systems of normal listeners, normal listeners with simulated hearing loss and hearing-impaired listeners. The carrier of the modulated signal was a noise having a long-term spectrum approximating that of speech; the modulation frequencies were selected to span the range of the modulation spectrum of speech (2.5–20 Hz). An overall level representative of conversational speech was also employed.

METHODS

Subjects

Two groups of subjects were included in the present study. One group consisted of 10 young adults (17–32 years of age) with normal hearing and normal middle-ear function [tymanograms of normal shape and amplitude and acoustic reflexes present at 105 dB HL (ANSI, 1970) from 500–4000 Hz]. All 10 subjects had PMTFs measured in quiet and in a background of high-pass noise designed to simulate the hearing loss of the impaired listeners. The same 10 subjects served as both the group of normal listeners and the group of normal listeners with simulated hearing loss.

The other group of listeners consisted of 5 adults ranging in age from 22 to 67 years. Their ages and their audiometric data for the test ear are summarized in Table 1. Each subject had a bilaterally symmetrical sloping sensorineural hearing loss and normal middle-ear function (normal tymanograms and acoustic reflexes present at 105 dB HL at 1000 Hz).

Stimulus Generation

The carrier of modulation was speech noise generated by a Grason-Stadler 901B noise source. The one-third octave-band spectrum of this noise measured in an NBS-9A 6-cm³ coupler showed an increase of 7 dB from 100 through 1000 Hz and a decrease of about 8 dB/octave above 1000 Hz. The carrier was modulated by cosinusoidal modulators with dc-offset that were digitally generated to simulate the range of the modulation spectrum of speech (2.5–20 Hz). An overall level representative of conversational speech was also employed.

All stimuli were presented monaurally through a TDH-49 earphone mounted in a circumaural cushion (Grason-Stadler 001). All testing was performed in a double-walled room having ambient noise levels below those specified by ANSI S3.1-1977 for earphone testing.

The noise used to simulate the high-frequency hearing loss of the impaired subjects was produced by a second noise generator (GenRad 1390B). This noise was routed through a bandpass filter having cutoff frequencies of 2000 and 10 000 Hz and rejection rates of 24 dB/octave. The overall level of this noise was adjusted to produce an estimated amount of masking at 4000 Hz generally approximating the average hearing loss of the impaired subjects at this same frequency. The lower cutoff frequency of the bandpass noise was then adjusted to produce an amount of masking at 1400 Hz that generally represented the average hearing loss of the impaired listeners at that frequency. No attempt was made to simulate the details of the audiometric configuration of the impaired listeners; only a general approximation was desired. The model of Humes et al. (1988) compensates for differences between the simulated and actual hearing loss. These preliminary adjustments of the masking noise were made with a separate group of 3 normal listeners and then applied to all of the normal subjects in this experiment. All stimuli were presented monaurally through a TDH-49 earphone mounted in a circumaural cushion (Grason-Stadler 001). All testing was performed in a double-walled room having ambient noise levels below those specified by ANSI S3.1-1977 for earphone testing.

Procedure

An adaptive two-interval forced-choice (2IFC) method was used to measure the threshold for the tone burst in
the presence of the modulated noise. The level of the tone burst was adjusted adaptively to estimate the 70.7% point of the psychometric function (Levitt, 1971). A total of 15 reversals in signal level comprised a single run. The first 2 reversals and the last reversal were discarded with the remaining 12 averaged to yield a single threshold estimate. Step size of attenuation was 10 dB for the first 2 reversals and 2 dB thereafter. Three thresholds were measured for a single condition prior to proceeding to the next condition.

The subjects were instructed to listen for a tone burst embedded in a modulated noise at one of two intervals marked by a light. The intervals were marked approximately 400 and 800 ms into the 1200-ms modulated noise sample. The interval locations varied slightly with frequency to assure that the presentation of the tone burst coincided with a maximum or minimum in the envelope of the modulated noise. Thus, the burst coincided with the acoustic "peak" or "trough" of the modulated noise. For the low modulation frequencies examined here, the maxima and minima in the acoustic noise waveforms correspond to the points of maximum and minimum masking (Scott, 1986).

Quiet thresholds for all subjects were also measured for the tone burst using the 2IFC procedure. Measurements were repeated for the normal listeners in the presence of a bandpass noise adjusted, in the manner described previously, to approximate the average hearing loss of the impaired listeners. The normal listeners were always tested first without the bandpass-noise background.

RESULTS AND DISCUSSION

Means and standard deviations for the quiet thresholds for the tone bursts and for the masked thresholds representing the simulated hearing loss are provided in Table 2. The listeners with simulated hearing loss had masked thresholds that were slightly lower than those of the listeners with actual hearing loss at 500 and 4000 Hz, whereas the simulated loss exceeded the actual loss at 1400 Hz by approximately 9 dB.

The mean masked thresholds for the 4.6-ms probe signal are presented in Figure 1; Table 3 contains the standard deviations for these data. The data from the normal listeners appear in the top panel of the figure, the data from normal listeners with simulated hearing loss in the middle panel, and the data from the hearing-impaired listeners in the bottom panel. Circles represent masked

![Figure 1](image-url)
TABLE 3. Standard deviations in dB for the data shown in Figure 1. Bold entries in table are for the tone burst at the peak of the modulated noise, whereas other entries are for the tone burst at the trough of the modulated noise.

<table>
<thead>
<tr>
<th>Group</th>
<th>Signal frequency in Hz</th>
<th>Modulation frequency in Hz</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>2.5</td>
</tr>
<tr>
<td>Normal hearing—Quiet</td>
<td>500</td>
<td>3.7</td>
</tr>
<tr>
<td></td>
<td>1400</td>
<td>2.8</td>
</tr>
<tr>
<td></td>
<td>4000</td>
<td>4.2</td>
</tr>
<tr>
<td>Normal hearing—Simulated Loss</td>
<td>500</td>
<td>6.2</td>
</tr>
<tr>
<td></td>
<td>1400</td>
<td>4.9</td>
</tr>
<tr>
<td></td>
<td>4000</td>
<td>2.8</td>
</tr>
<tr>
<td>Hearing impaired</td>
<td>500</td>
<td>1.2</td>
</tr>
<tr>
<td></td>
<td>1400</td>
<td>8.6</td>
</tr>
<tr>
<td></td>
<td>4000</td>
<td>13.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>12.1</td>
</tr>
</tbody>
</table>

thresholds for the probe signal temporally positioned at the acoustic maximum (peak) of the modulated-noise envelope and the squares represent corresponding thresholds for the probe signal in the acoustic minimum (trough) of the modulated masker. Three sets of masked thresholds appear in each panel, one for each probe frequency (500, 1400, and 4000 Hz).

Several features of the data in the top panel for the normal-hearing subjects listening in quiet are noteworthy. First, the masked threshold for the probe signal at the peak of the modulator envelope (circles) does not vary with modulation frequency. The mean thresholds averaged across modulation frequency for each probe frequency are 81.8, 81.5 and 65.7 dB SPL at 500, 1400, and 4000 Hz, respectively. Second, the masked threshold for the probe in the trough in the modulator (squares) increases as modulation frequency increases. Recall that the noise was 100% amplitude modulated at all modulation frequencies. As modulation frequency increases, the time separation between successive peaks in the modulator is decreased. Thus, the probe signal, centered between successive peaks of the modulator, is subject to greater amounts of forward and backward masking from the immediately preceding and following peaks of the modulated masker as the modulation frequency is increased (Fastl, 1976). The increased temporal masking of the probe signal with increased modulation frequency results in the upward sloping functions for the probe-in-trough conditions. Finally, the thresholds for both temporal probe locations, peak and trough, are lower at 4000 Hz than at the other two probe frequencies. This follows directly from the lower amount of high-frequency acoustic energy in the speech-shaped carrier used in this study.

The middle and bottom panels of Figure 1 display data for the other two groups of listeners. For both of these groups, the data at probe frequencies of 500 and 1400 Hz are similar to those in the top panel in that the thresholds for the probe signal located in the acoustic maximum of the modulated masker do not vary with modulation frequency, whereas those obtained in the trough of the modulated noise increase with modulation frequency. This pattern can again be explained by increasing amounts of forward and backward masking for the probe-in-trough conditions as the modulation frequency increases. At 4000 Hz, the amount of hearing loss, either simulated (middle panel) or actual (bottom panel), exceeded the masked thresholds for the normal listeners in quiet, so “masked” thresholds didn’t vary with modulation frequency or temporal location of the probe signal.

The solid lines in the lower two panels in Figure 1 represent the predictions made by the modified power-law model of Humes et al. (1988). This model, as applied here, is in the following form: $Y = X_1^P + X_2^P - X_Q^P$, where $Y$ is the predicted signal intensity (in W/cm²) at masked threshold for the modulated-noise masker, $X_1$ is the signal intensity at masked threshold for the same modulated-noise masker in the normal ears, $X_2$ is the signal intensity at masked threshold for the simulated hearing loss or the quiet threshold of the subject with actual hearing loss, and $X_Q$ is the signal intensity at quiet threshold for the normal listeners. The exponent ($P$) in the power-law model was fixed at 0.3 as in previous applications to the data from hearing-impaired listeners by Humes et al. (1988). These predictions were made by combining the masked thresholds, $X_1$, in the top panel of Figure 1 with either the mean masked threshold representing the simulated hearing loss (middle panel) or the mean quiet threshold of the impaired subjects (bottom panel), $X_2$. The solid lines without symbols provide a reasonable description of the data at 500 and 4000 Hz for both subject groups. At 1400 Hz, the model overestimates...
the masked threshold for the probe-in-peak condition (circles) by about 5 dB for both groups of subjects. Still, using a fixed exponent of 0.3, 94% of the variance in the data can be accounted for by the model at probe frequencies of 500 and 1400 Hz. At 4000 Hz, where neither predicted or observed thresholds vary by more than 5 dB across all modulation frequencies, only 41% of the variance is accounted for by the model with $P = 0.3$. When evaluated across all probe frequencies, modulation frequencies, and subject groups, the model accounts for 86% of the variance in the data for $P = 0.3$.

In general, the data in Figure 1 suggest that the auditory system of normal listeners can preserve envelope fluctuations of up to 30–40 dB at lower modulation frequencies (2.5 Hz) and only up to 10–15 dB at higher modulation frequencies (20 Hz) within the modulation spectrum of speech. Thus, acoustic fluctuations of 30 dB in the envelope of speech will be encoded as comparable fluctuations in perceptual magnitude only for slowly occurring changes. Other investigations with normal listeners have also reported a reduction in psychoacoustic modulation depth as modulation frequency increased over a similar range. Two previous studies, for example, observed a reduction in psychoacoustic modulation depth of 6–8 dB as the modulation frequency increased from 10 to 20 Hz (Fastl, 1976; Festen & Plomp, 1981). In comparison, results from the present study for a comparable signal frequency (1400 Hz) and the same range of modulation frequencies exhibit a 9-dB reduction in psychoacoustic modulation depth. Scott (1986), using modulation frequencies from 2 to 256 Hz and signal frequencies of 1000 and 4000 Hz, obtained results virtually identical to those reported here.

Scott (1986) also obtained psychophysical modulation transfer functions, for the same subjects and conditions, using two other methods: a masking method using long-duration signals (e.g., Weber, 1977) and a modulation-detection method (e.g., Viemeister, 1979). When plotted as modulation transfer functions, the two masking-based measures revealed a decrease in modulation preservation within the auditory system for modulation frequencies greater than 8–16 Hz. The modulation-detection paradigm, however, indicated a flat modulation transfer function out to 75–100 Hz. It was concluded by Scott (1986) that the various psychophysical methods available to measure modulation transfer functions are not equivalent. Psychophysical modulation transfer functions for the preservation of modulation depth for a 100% amplitude-modulated stimulus do not appear to be equivalent to those for the detection of the minimum amount of modulation.

In the present study, the noise-masked normal listeners with simulated hearing loss and the hearing-impaired listeners both have a reduced range over which envelope fluctuations can be encoded. The narrower range in both groups, however, is consistent with that predicted by the model of Humes et al. (1988). In generating predictions, this model takes into consideration only the elevated thresholds of the listeners with actual or simulated hearing loss and the accompanying loudness recruitment. No separate and independent alterations in a temporal-resolution mechanism are assumed. That is, the reduced range over which the envelope fluctuations are encoded can be described as a simple consequence of threshold elevation accompanied by recruitment. Whereas the model of Humes et al. also provided an accurate description of the results from Zwickler and Schorn (1982) for normal listeners with simulated losses, such was not the case for their data from hearing-impaired subjects. The model underestimated masked thresholds for the modulated masker by approximately 20 dB when applied to impaired listeners. A different paradigm was employed in that study, however, in which a long-duration signal was detected in a background of modulated noise (100% square-wave modulation with a 15.6-Hz modulation frequency). Perhaps differences in paradigm are responsible for the differences in outcome.

ACKNOWLEDGMENTS

This research was supported, in part, by grants from NIH, the Veterans Administration, and the Air Force Office of Scientific Research. The expert assistance of Dr. Christopher Ahlstrom in all phases of data collection and reduction is gratefully acknowledged.

REFERENCES


Received March 13, 1989
Accepted June 23, 1989

Requests for reprints should be sent to Larry E. Humes, Department of Speech and Hearing Sciences, Indiana University, Bloomington, IN 47405.