A COMPARISON OF THE EFFECTS OF FILTERING AND SENSORINEURAL HEARING LOSS ON PATTERNS OF CONSONANT CONFUSIONS

MARILYN D. WANG

University of Maryland Baltimore County, Baltimore

CHARLOTTE M. REED *and* ROBERT C. BILGER

University of Pittsburgh, Pennsylvania

It has been found that listeners with sensorineural hearing loss who show similar patterns of consonant confusions also tend to have similar audiometric profiles. The present study determined whether normal listeners, presented with filtered speech, would produce consonant confusions similar to those previously reported for the hearing-impaired listener. Consonant confusion matrices were obtained from eight normal-hearing subjects for four sets of CV and VC nonsense syllables presented under six high-pass and six low-pass filtering conditions. Patterns of consonant confusion for each condition were described using phonological features in a sequential information analysis. Severe low-pass filtering produced consonant confusions comparable to those of listeners with high-frequency hearing loss. Severe high-pass filtering gave a result comparable to that of patients with flat or rising audiograms. And, mild filtering resulted in confusion patterns comparable to those of listeners with essentially normal hearing. An explanation in terms of the spectrum, the level of speech, and the configuration of the individual listener's audiogram is given.

In a study of the consonant confusions of listeners with sensorineural hearing loss, Bilger and Wang (1976) found that patients with similar patterns of perceptual confusions tended also to have similar audiometric configurations From a feature analysis of the perceptual confusions, three groups of listeners were identified: listeners with moderate or severe pure-tone losses at all frequencies; listeners with normal hearing or mild pure-tone losses at all frequencies; and listeners with predominantly high-frequency losses or steeply sloping audiograms.

These findings suggest that for listeners with sensorineural hearing loss, patterns of consonant confusions can be predicted from knowledge of the pure-tone audiogram. However, if it is simply the loss or attenuation of specific frequency regions that results in particular confusion patterns, then it should also be possible to generate comparable patterns in normal listeners by filtering the speech signal.

There is already some evidence that when normal listeners are presented with filtered speech they produce confusions similar to those produced by hearing-impaired listeners. Using a multiple-choice word-recognition test, Owens, Benedict, and Schubert (1972) compared phonemic errors of listeners with sensorineural hearing loss above 500 Hz, with those of normal listeners presented with 780-Hz low-pass speech. They found that persons with similar pure-tone losses, but different etiologies, made the same kinds of errors, and that listeners with sharply sloping audiograms made errors similar to those made by the normal subjects with the filtered speech. Similarly, Sher and Owens (1974) compared a group of patients with high-frequency losses above 2000 Hz and normal subjects listening to 2040-Hz low-pass speech. They reported no difference between the groups in overall scores, probabilities of error for individual phonemes, or patterns of phonemic confusions. They also noted that the similarity between normal and hearing-impaired listeners was greatest when the rejection rate of the filter most closely approximated the slope of the hearing loss.

The present experiment extended this type of comparison by determining whether normal listeners presented with filtered speech would produce consonant confusions similar to those of the three groups of hearing-impaired listeners described by Bilger and Wang (1976). Both high-pass and low-pass conditions were included. We predicted that with mild filtering the performance of the normal subjects would approximate that of the normal and mild, flat-loss patients. As the low-pass cutoff was lowered, however, we expected that the consonant confusions would become similar to those of the high-frequency-loss group. Although Owens et al. (1972) and Sher and Owens (1974) did not use high-pass conditions, they were included here for two reasons. First, three of the subjects in the moderate-to-severe flat-loss group showed slightly better pure-tone thresholds at higher frequencies than at lower frequencies. Second, two of the characteristics of that group's performance were good perception of sibilance and relatively poor perception of nasality. Since sibilant sounds are distinguished by high-frequency cues, and since nasality is cued by low-frequency information, we predicted that high-pass conditions would produce performance similar to that of the "flat"-loss group.

METHOD

Subjects

The subjects were eight undergraduate students with normal hearing. They were divided into two listening groups of four subjects each. They were paid $2.00 per hour for their participation.

Stimuli

The stimuli for the experiment were recorded CV and VC nonsense syllables, assembled into four 48-syllable sets. The consonants in each set are given in

Table 1. Two sets, CV-1 and CV-2, contained only CV syllables; the other sets, VC-1 and VC-2, contained only VC syllables. In each syllable set, 16 consonants were combined with the three vowels /i, a, u/.

TABLE 1. Composition of four syllable sets.

Set	*Consonants*
CV-1	/p, t, k, b, d, g, f, θ, s, ʃ, v, ð, z, ʒ, tʃ, dʒ/
VC-1	/p, t, k, b, d, g, f, θ, s, ʃ, v, ð, z, ʒ, tʃ, dʒ/
CV-2	/p, b, t, d, l, r, f, s, v, z, m, n, h, hʷ, w, j/
VC-2	/p, b, g, ŋ, m, n, f, θ, s, ʃ, v, ð, z, ʒ, tʃ, dʒ/

Experimental Conditions

Twelve filter conditions were included. The wide-band condition was an 80–5600-Hz passband. For the high-pass conditions, the upper cutoff remained at 5600 Hz, while the lower cutoff took on values of 355, 710, 1400, 2000, 2800, and 4000 Hz. For the low-pass conditions, the lower cutoff remained at 80 Hz, while the upper cutoff took on values of 2800, 1400, 1000, 710, and 500 Hz. These conditions were selected to provide a range of performance and choices, and were guided by results presented by Hirsh, Reynolds, and Joseph (1954) for high- and low-pass filtering of nonsense syllables.

A single test list consisted of 96 randomly ordered items from one syllable set (two occurrences of each of the 48 syllables). For one replication of the experiment, each subject was presented with two test lists from each syllable set under each of the 12 filter conditions. The order of the filter conditions was randomized, and the order of syllable sets within conditions was also randomized. The first group of subjects received the original randomization, and the second group received the randomization in reverse order. Three complete replications of the experiment were performed. Thus, each subject provided 576 observations under each filter condition for each syllable set.

Procedure

The syllables were recorded on the addressable drum of a Cognitronics Speechmaker. The output of the Speechmaker was led through an electronic switch with a 1-msec rise time. The signal was then passed through two filters, each with a rejection rate of 24 dB/oct, for a total rejection rate of 48 dB/oct. The signal was then amplified, attenuated, and passed through a passive four-way splitter into calibrated TDH-49 earphones in four sound-treated booths. Each time a filter setting was changed, the new setting was adjusted using a voltage measurement of the half-power points.

The level of the wide-band condition was set at 95-dB SPL re a 1000-Hz calibration tone. A constant spectrum level of speech was then used for all filter conditions. The overall loudness level was thus a function of the band width of the filtered speech. For the high-pass conditions, there was a maximum 5-dB difference between the wide-band condition and the narrowest filter condition; for the low-pass conditions, this maximum difference was 13 dB. These differences, however, are partially accounted for by the spectral characteristics of speech.

Subjects were seated in individual sound-treated rooms, in front of a response console that contained a 4 × 4 array of response buttons with orthographic representations of the 16 consonants in the given syllable set. A list of cue words for the orthographic symbols was provided in each room. A trial consisted of a 500-msec warning interval, followed by a 511-msec observation interval during which a syllable was presented. Subjects responded by pressing one of the 16 response buttons. Correct-answer feedback was provided after every trial.

On a given day, subjects listened to eight test lists (two from each syllable set) under a single filter condition. Since there were three replications of the experiment, each subject provided a total of 576 observations per condition. The experiment took 10 weeks to complete, with subjects working two hours per day, four days per week. Included in this time was a week of practice listening to the four syllable sets under selected high- and low-pass filter settings.

RESULTS

Overall performance on the four syllable sets was evaluated by calculating the mean percentage of correct responses across listeners and filter conditions. The means were as follows: CV-1, 67.5%; VC-1, 80.0%; CV-2, 77.7%; and VC-2, 79.3%. Performance was significantly better on the VC sets than on the CV sets ($F = 118.22$, $df = 1{,}21$), and was also significantly better on the CV-2 set than on the CV-1 set ($F = 122.05$, $df = 1{,}21$). In our previous studies using these stimuli, the CV-1 set has been shown to be most difficult (Wang and Bilger, 1973; Bilger and Wang, 1976). Differences among the remaining sets have been somewhat less consistent, but filtering seems to degrade relative performance on the CV-2 set more than other types of distortion.

Performance on the syllable sets as a function of filter conditions is shown in the four panels of Figure 1. The wide-band condition (80–5600 Hz) is plotted in both the high-pass and low-pass functions. The crossover point of these two functions varies from 1750 to 2000 Hz for the different syllable sets. These crossover frequencies are in close agreement both with the 1900-Hz crossover reported by French and Steinberg (1947) for CVC syllables, and with the 1700-Hz crossover reported by Hirsh et al. (1954) for CV, VC, and CVC nonsense syllables. They are slightly higher than the 1550-Hz point reported by Miller and Nicely (1955) for CV syllables. For all syllable sets, low-pass filtering caused a greater decrement in performance than did high-pass filtering, although the difference was smaller for the CV-2 set than for the

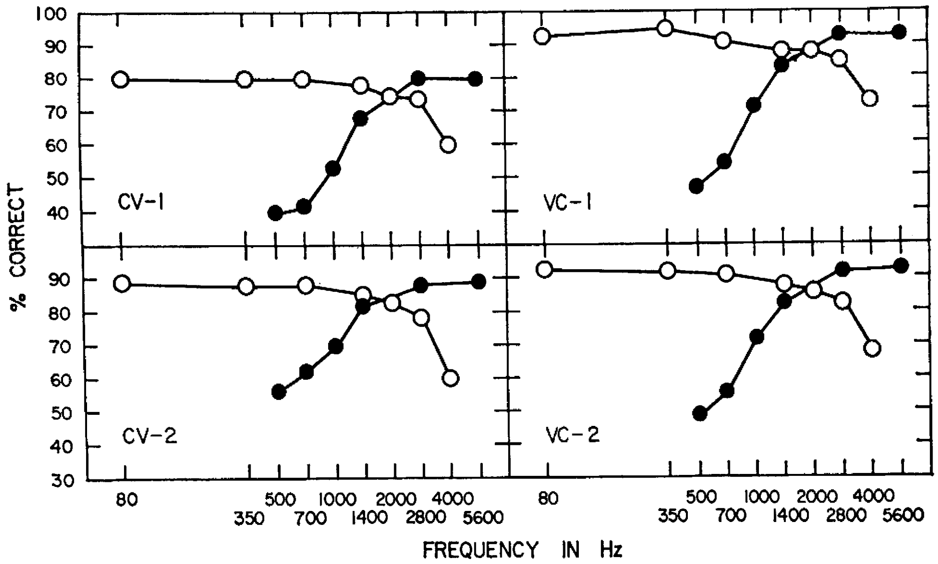

Figure 1. Performance on four syllable sets as a function of filter cutoff. Filled symbols represent low-pass conditions. Open symbols represent high-pass conditions. The wide-band condition is plotted in both functions. Each point is based on approximately 4608 observations.

others. In fact, for low-pass filtering at 710 Hz and 500 Hz, performance on the CV-2 set was significantly better than that on the other three sets ($F = 17.76$ and $F = 19.15$, $df = 1{,}21$, $p < 0.05$).

Uncertainty Analysis of Perceptual Confusions

The data of primary interest in the present study were the patterns of consonant confusions obtained under the different filtering conditions. In previous studies we have shown that confusion patterns can be conveniently summarized using a type of multivariate uncertainty analysis that focuses on the transmitted information associated with a given stimulus-response confusion matrix, and that identifies the contributions of various phonological features to the transmitted information. The procedure, which we have called sequential information analysis (SINFA), sequentially identifies features with a high proportion of transmitted information. At each step, however, the contributions of features previously identified are partialed out. This makes it possible to control for feature redundancy and to identify the independent contributions of several features to the total information transmitted. A more detailed description of the procedure is given in Wang and Bilger (1973).

Consonant confusion matrices were constructed from the pooled responses of all subjects for each condition and syllable set. (Copies of these matrices are available from the authors.) A set of familiar descriptive features was used as a basis for the sequential information analysis (see Table 2). A summary of

TABLE 2. Features specified for the sequential information analysis. (See Bilger and Wang, 1976)

Phoneme	*Voc*	*Cons*	*High*	*Low*	*Back*	*Cor*	*Ant*	*Voi*	*Nas*	*Cont*	*Str*	*Rnd*	*Fric*	*Dur*	*Pl*	*Sib*
p	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0
t	0	1	0	0	0	1	1	0	0	0	0	0	0	0	1	0
k	0	1	1	0	1	0	0	0	0	0	0	0	0	0	4	0
b	0	1	0	0	0	0	1	1	0	0	0	0	0	0	0	0
d	0	1	0	0	0	1	1	1	0	0	0	0	0	0	1	0
g	0	1	1	0	1	0	0	1	0	0	0	0	0	0	4	0
f	0	1	0	0	0	0	1	0	0	1	1	0	1	0	0	0
θ	0	1	0	0	0	1	1	0	0	1	0	0	1	0	1	0
s	0	1	0	0	0	1	1	0	0	1	1	0	1	1	2	1
ʃ	0	1	1	0	0	1	0	0	0	1	1	0	1	1	3	1
v	0	1	0	0	0	0	1	1	0	1	1	0	1	0	0	0
ð	0	1	0	0	0	1	1	1	0	1	0	0	1	0	1	0
z	0	1	0	0	0	1	1	1	0	1	1	0	1	1	2	1
ʒ	0	1	1	0	0	1	0	1	0	1	1	0	1	1	3	1
tʃ	0	1	1	0	0	1	0	0	0	0	1	0	1	0	3	1
dʒ	0	1	1	0	0	1	0	1	0	0	1	0	1	0	3	1
m	0	1	0	0	0	0	1	1	1	0	0	0	0	0	0	0
n	0	1	0	0	0	1	1	1	1	0	0	0	0	0	1	0
ŋ	0	1	1	0	1	0	0	1	1	0	0	0	0	0	4	0
r	1	1	0	0	0	1	0	1	0	1	0	0	0	0	1	0
l	1	1	0	0	0	1	1	1	0	1	0	0	0	0	2	0
w	0	0	1	0	1	0	0	1	0	1	0	1	0	0	0	0
h	0	0	0	1	0	0	0	0	0	1	0	0	1	0	4	0
hʷ	0	0	0	1	0	0	0	0	0	1	0	1	1	0	4	0
j	0	0	1	0	0	0	0	1	0	1	0	0	0	0	3	0

Table 3. Sequential information analyses for CV-1 syllable set. The table entries give the (conditional) information transmitted for different features; numbers in parentheses give the iteration in which the feature had the highest proportion of information transmitted. Tr. inf. = transmitted information.

				Filter Condition (Hz)			
		80-5600	*80-2800*	*80-1400*	*80-1000*	*80-710*	*80-500*
Feature	*Tr. inf.:*	*3.205*	*3.236*	*2.573*	*1.965*	*1.579*	*1.366*
Voice		0.906 (2)	0.925 (2)	0.885 (1)	0.806 (1)	0.735 (1)	0.584 (1)
Frication		–	0.285 (4)	–	0.129 (4)	0.328 (2)	0.297 (2)
Sibilant		0.873 (1)	0.889 (1)	0.226 (4)	0.441 (2)	0.051 (4)	–
Strident		–	–	–	–	–	–
Duration		0.143 (4)	–	0.103 (5)	–	–	0.066 (4)
Place		–	0.198 (6)	0.100 (6)	–	–	–
High/Anter		0.694 (3)	0.695 (3)	0.661 (2)	0.279 (3)	0.226 (3)	0.189 (3)
Back		–	–	0.227 (3)	–	–	–
Coronal		0.193 (6)	–	–	0.049 (6)	0.021 (6)	0.031 (5)
Continuant		0.282 (5)	0.131 (5)	0.181 (6)	0.052 (5)	0.021 (5)	–
Sum		3.091	3.123	2.383	1.756	1.382	1.167
% Tr. inf.		96.4	96.5	92.6	89.4	87.5	85.4

		355-5600	*710-5600*	*1400-5600*	*2000-5600*	*2800-5600*	*4000-5600*
Feature	*Tr. inf.:*	*3.203*	*3.144*	*3.040*	*2.929*	*2.778*	*2.190*
Voice		0.895 (2)	0.855 (2)	0.814 (2)	0.743 (3)	0.718 (2)	0.614 (2)
Frication		–	–	–	–	–	–
Sibilant		0.884 (1)	0.855 (1)	0.833 (1)	0.844 (1)	0.862 (1)	0.732 (1)
Strident		–	–	–	–	–	–
Duration		0.147 (4)	0.149 (4)	0.149 (4)	0.143 (4)	0.124 (4)	0.188 (3)
Place		–	–	–	–	–	–
High/Anter		0.684 (3)	0.664 (3)	0.640 (3)	0.590 (2)	0.572 (3)	0.223 (4)
Back		–	–	–	–	–	–
Coronal		0.198 (6)	0.216 (6)	0.198 (6)	0.185 (6)	0.156 (6)	0.084 (6)
Continuant		0.277 (5)	0.274 (5)	0.244 (5)	0.239 (5)	0.219 (5)	0.172 (5)
Sum		3.085	3.013	2.878	2.744	2.651	2.013
% Tr. inf.		96.3	95.8	94.7	93.7	95.4	91.9

Table 4. Sequential information analyses for VC-1 syllable set. The table entries give the (conditional) information transmitted for different features; numbers in parentheses give the iteration in which the feature had the highest proportion of information transmitted. Tr. inf. = Transmitted information.

		Filter Condition (Hz)					
		80-5600	*80-2800*	*80-1400*	*80-1000*	*80-710*	*80-500*
Feature	*Tr. inf.:*	*3.579*	*3.605*	*3.186*	*2.626*	*2.044*	*1.795*
Voice		0.867 (4)	0.891 (4)	0.857 (3)	0.818 (1)	0.796 (1)	0.784 (1)
Frication		0.440 (3)	–	0.825 (1)	0.738 (2)	0.182 (3)	0.455 (2)
Sibilant		0.913 (1)	0.918 (1)	0.091 (6)	0.072 (6)	0.076 (5)	–
Strident		–	–	–	–	–	–
Duration		–	0.316 (2)	–	–	–	0.050 (5)
Place		–	–	–	–	–	–
High/Anter		0.720 (2)	0.521 (5)	0.581 (4)	0.326 (4)	–	–
Back		–	–	–	–	0.133 (4)	0.084 (4)
Coronal		0.353 (6)	0.347 (6)	0.315 (5)	0.201 (5)	0.080 (6)	0.033 (5)
Continuant		0.208 (5)	0.551 (3)	0.405 (2)	0.306 (3)	0.599 (2)	0.218 (3)
Sum		3.501	3.544	3.074	2.461	1.866	1.624
% Tr. inf.		97.8	98.3	96.5	93.7	91.3	90.5
		355-5600	*710-5600*	*1400-5600*	*2000-5600*	*2800-5600*	*4000-5600*
Feature	*Tr. inf.:*	*3.628*	*3.551*	*3.365*	*3.399*	*3.221*	*2.676*
Voice		0.899 (3)	0.843 (4)	0.727 (5)	0.744 (5)	0.677 (4)	0.578 (4)
Frication		0.423 (4)	–	0.384 (4)	–	–	–
Sibilant		0.923 (1)	0.927 (1)	0.904 (1)	0.921 (1)	0.903 (1)	0.872 (1)
Strident		–	–	–	–	–	–
Duration		–	0.322 (2)	0.316 (2)	0.208 (3)	0.217 (3)	0.269 (2)
Place		–	–	–	–	0.281 (6)	0.165 (6)
High/Anter		0.716 (2)	0.499 (5)	0.619 (3)	0.745 (2)	0.724 (2)	0.424 (3)
Back		–	–	–	–	–	–
Coronal		0.364 (6)	0.336 (6)	0.325 (6)	0.316 (6)	–	–
Continuant		0.213 (5)	0.545 (3)	–	0.379 (4)	0.326 (5)	0.259 (5)
Sum		3.538	3.472	3.275	3.313	3.128	2.567
% Tr. inf.		97.5	97.8	97.3	97.5	97.1	95.9

TABLE 5. Sequential information analyses for CV-2 syllable set. The table entries give the (conditional) information transmitted for different features; numbers in parentheses give the iteration in which the feature had the highest proportion of information transmitted. Tr. inf. = Transmitted information.

		Filter Condition (Hz)					
		80-5600	*80-2800*	*80-1400*	*80-1000*	*80-710*	*80-500*
Feature	*Tr. inf.:*	*3.552*	*3.503*	*3.188*	*2.623*	*2.299*	*2.095*
Vocalic		0.431 (4)	0.524 (1)	0.474 (2)	0.368 (2)	0.183 (4)	0.172 (3)
Consonantal		–	0.291 (6)	0.546 (3)	0.453 (3)	0.205 (5)	0.284 (4)
High		–	–	–	–	–	–
Low		–	–	–	–	0.174 (3)	–
Back		–	–	–	0.055 (6)	0.037 (8)	–
Coronal		0.104 (7)	–	–	–	–	–
Anterior		0.398 (6)	0.102 (7)	0.443 (4)	0.282 (5)	0.129 (7)	0.159 (6)
Place		–	0.093 (9)	0.055 (8)	–	–	–
Voice		0.593 (8)	0.577 (8)	0.562 (6)	0.468 (4)	0.514 (2)	0.493 (2)
Nasal		0.471 (3)	0.429 (4)	0.490 (1)	0.449 (1)	0.432 (1)	0.436 (1)
Duration		0.244 (2)	0.248 (3)	–	–	–	–
Strident		0.118 (9)	0.125 (10)	0.095 (9)	0.066 (9)	0.058 (10)	0.152 (5)
Sibilant		0.795 (1)	0.733 (2)	0.192 (7)	0.151 (8)	0.178 (6)	0.026 (8)
Continuant		–	–	–	–	–	–
Round		0.344 (5)	0.325 (5)	0.201 (5)	0.051 (7)	0.036 (9)	0.040 (7)
Sum		3.498	3.447	3.058	2.343	1.946	1.762
% Tr. inf.		98.5	98.4	95.9	89.3	84.6	84.1

Table 5. (cont.)

Feature		355-5600	710-5600	1400-5600	2000-5600	2800-5600	4000-5600
	Tr. inf.:	*3.495*	*3.478*	*3.352*	*3.178*	*2.867*	*2.107*
Vocalic		0.457 (3)	0.445 (3)	0.199 (6)	0.144 (7)	0.306 (3)	–
Consonantal		–	–	0.389 (4)	0.527 (3)	–	–
High		–	–	–	–	–	0.054 (6)
Low		–	–	–	–	–	0.124 (4)
Back		–	–	–	–	0.057 (6)	–
Coronal		–	0.263 (4)	–	–	–	0.074 (7)
Anterior		0.509 (5)	0.424 (6)	0.241 (5)	0.222 (5)	0.451 (4)	0.140 (5)
Place		0.094 (8)	–	0.147 (8)	0.125 (8)	0.107 (7)	–
Voice		0.577 (7)	0.618 (7)	0.646 (7)	0.631 (6)	0.521 (5)	0.412 (3)
Nasal		0.416 (4)	0.117 (8)	0.082 (9)	0.087 (9)	0.040 (10)	0.026 (10)
Duration		0.249 (2)	0.535 (1)	0.248 (2)	0.525 (1)	0.527 (1)	0.206 (2)
Strident		0.117 (9)	0.123 (9)	0.101 (10)	0.110 (10)	0.120 (9)	0.063 (9)
Sibilant		0.797 (1)	0.507 (2)	0.794 (1)	0.502 (2)	0.489 (2)	0.681 (1)
Continuant		–	–	–	–	–	–
Round		0.209 (6)	0.367 (5)	0.424 (3)	0.213 (4)	0.047 (8)	0.044 (8)
Sum		3.425	3.399	3.271	3.086	2.665	1.824
% Tr. inf.		98.0	97.7	97.6	97.1	93.0	86.6

TABLE 6. Sequential information analyses for VC-2 syllable set. The table entries give the (conditional) information transmitted for different features; numbers in parentheses give the iteration in which the feature had the highest proportion of information transmitted. Tr. inf. = Transmitted information.

		Filter Condition (Hz)					
		80-5600	*80-2800*	*80-1400*	*80-1000*	*80-710*	*80-500*
Feature	*Tr. inf.:*	3.584	3.539	*3.103*	*2.682*	*2.153*	*1.831*
Voice		0.645 (4)	0.639 (5)	0.660 (4)	0.792 (1)	0.765 (1)	0.717 (1)
Nasal		0.479 (3)	0.477 (3)	0.596 (1)	0.436 (2)	0.386 (2)	0.319 (2)
Frication		–	–	0.536 (2)	0.469 (3)	0.140 (4)	0.137 (4)
Sibilant		0.922 (1)	0.906 (1)	0.146 (6)	0.083 (7)	–	0.035 (6)
Strident		–	–	–	–	–	–
Duration		0.212 (5)	0.216 (4)	–	–	0.061 (6)	–
Place		–	–	–	–	–	–
High/Anter		0.738 (2)	0.726 (2)	0.526 (5)	0.181 (6)	–	–
Back		–	–	–	0.137 (5)	0.109 (5)	0.072 (5)
Coronal		0.222 (7)	0.206 (7)	0.122 (7)	0.086 (8)	0.044 (7)	0.018 (7)
Continuant		0.277 (6)	0.274 (6)	0.391 (3)	0.329 (4)	0.476 (3)	0.368 (3)
Sum		3.495	3.444	2.977	2.513	1.981	1.666
% Tr. inf.		97.5	97.3	95.9	93.7	92.0	91.0

		355-5600	*710-5600*	*1400-5600*	*2000-5600*	*2800-5600*	*4000-5600*
Feature	*Tr. inf.:*	*3.561*	3.478	*3.369*	*3.271*	3.085	2.486
Voice		0.765 (3)	0.620 (5)	0.567 (5)	0.648 (4)	0.615 (4)	0.539 (3)
Nasal		0.247 (5)	0.490 (3)	0.420 (4)	0.296 (5)	0.262 (5)	0.121 (6)
Frication		0.405 (4)	–	–	–	–	–
Sibilant		0.916 (1)	0.908 (1)	0.928 (1)	0.909 (1)	0.896 (1)	0.867 (1)
Strident		–	–	–	–	–	–
Duration		–	0.315 (2)	0.320 (2)	0.203 (3)	0.206 (3)	0.289 (2)
Place		–	–	–	–	–	0.060 (7)
High/Anter		0.724 (2)	0.602 (4)	0.607 (3)	0.699 (2)	0.689 (2)	0.351 (4)
Back		–	–	–	–	–	–
Coronal		0.219 (7)	0.205 (7)	0.192 (7)	0.176 (7)	0.138 (7)	–
Continuant		0.206 (6)	0.252 (6)	0.244 (6)	0.239 (6)	0.181 (6)	0.164 (5)
Sum		3.482	3.392	3.278	3.170	2.987	2.391
% Tr. inf.		97.8	97.5	97.3	96.9	96.8	96.2

those analyses is presented in Tables 3-6, which present the analysis for the CV-1, VC-1, CV-2, and VC-2 syllable sets, respectively. The top half of each table gives results for low-pass filter conditions. The bottom half gives results for high-pass conditions. Entries in the tables give the transmitted information for the confusion matrix, the transmitted information for the best feature in each iteration, and the iteration number. For all iterations after the first, the values represent conditional transmitted information, that is, transmitted information independent of the information transmitted by features identified in previous iterations. The percentage of transmitted information accounted for by the features identified by SINFA is shown at the bottom of each column in Tables 3-6. Since the lowest value is 84.1% and the majority are over 90%, the analyses provide an acceptable descriptive summary of the subjects' perceptual confusions.

If the first two columns of each table are arbitrarily designed as "mild" distortion conditions, we can identify six features that tend to be well perceived in all syllable sets where they are distinctive. The best perceived feature in 14 of the 16 analyses is sibilance; in the remaining two analyses sibilance ranks second. Although the ranking of the other strong features changes more across conditions and syllable sets, voice, high/anterior, and either frication or duration are usually among the first four or five features identified. Nasality, which is only distinctive in the CV-2 and VC-2 sets, is identified early in seven of the eight analyses. Vocalic, which is only distinctive in the CV-2 set, is also a relatively important feature.

Low-pass filtering produces systematic changes in the importance of different features. When the low-pass cutoff is lowered from 2800 to 1400 Hz, sibilance quickly loses its first-place ranking and drops to a ranking of sixth or worse for all sets except CV-1, where its ranking is less affected. Except for the CV-1 set, high and anterior also drop noticeably when the cutoff is lowered to 1400 Hz. Voice and nasality, however, become increasingly important as the low-pass cutoff is lowered. Frication and the related feature, duration, show inconsistent changes with low-pass filtering. Duration seems to be the weaker of the two features since it is identified less often. However, except for the CV-2 set, one or the other of these two features is still identified in the first four iterations. High-pass filtering produces less consistent changes in feature recognition, a result also noted by Miller and Nicely (1955). Sibilance remains high in the ranking of features under all high-pass conditions. High/anterior also does not change across conditions, although it ranks much lower than sibilance. Duration, rather than frication, ranks second or third in all sets except CV-1, where it is fourth. Voice is a strong feature, especially in the CV-1 set. Nasality, however, drops out completely in the CV-2 set and drops rank considerably in the VC-2 set.

For those who might be interested, we have analyzed the confusion matrixes reported by Miller and Nicely (1955) for high-pass and low-pass conditions, using SINFA. The results of these analyses are summarized in Table 7. Detailed

TABLE 7. Sequential information analyses for data reported by Miller and Nicely (1955). Tr. inf. = Transmitted information.

		Filter Condition (Hz)					
		200-5000	*200-2500*	*200-1200*	*200-600*	*200-400*	*200-300*
Feature	*Tr. inf.:*	*3.186*	*2.827*	*2.378*	*2.185*	*1.685*	*1.155*
High/Anter		0.567 (4)	0.425 (4)	–	–	–	–
Fric/Cont		0.661 (3)	0.616 (3)	0.535 (3)	0.535 (3)	0.336 (3)	0.128 (3)
Nasal		0.521 (1)	0.543 (1)	0.523 (1)	0.520 (1)	0.456 (1)	0.371 (1)
Voice		0.743 (2)	0.746 (2)	0.702 (2)	0.721 (2)	0.591 (2)	0.548 (2)
Place		0.427 (5)	0.303 (5)	0.318 (5)	0.149 (5)	0.088 (5)	0.035 (4)
Dur/Sib		0.101 (6)	0.032 (6)	0.162 (4)	0.146 (4)	0.095 (4)	–
Sum		3.020	2.665	2.240	2.071	1.566	1.082
% Tr. inf.		94.8	94.3	94.2	94.8	92.9	93.7
		1000-5000	*2000-5000*	*2500-5000*	*3000-5000*	*4500-5000*	
Feature	*Tr. inf.:*	*2.652*	*1.582*	*1.052*	*0.623*	*0.454*	
High/Anter		0.409 (4)	0.279 (2)	0.209 (2)	0.124 (2)	0.094 (2)	
Fric/Cont		0.229 (6)	0.103 (6)	0.055 (6)	0.021 (5)	0.008 (4)	
Nasal		0.350 (1)	0.082 (5)	0.036 (4)	0.008 (6)	–	
Voice		0.545 (3)	0.285 (3)	0.143 (3)	0.060 (3)	0.028 (3)	
Place		0.326 (5)	0.137 (4)	0.057 (5)	0.023 (4)	0.012 (5)	
Dur/Sib		0.472 (2)	0.425 (1)	0.348 (1)	0.235 (1)	0.192 (1)	
Sum		2.331	1.311	0.848	0.471	0.334	
% Tr. inf.		87.9	82.9	80.6	75.6	73.6	

comparisons of our data with theirs are difficult because the two studies differ in several respects:

(1) The syllable set used by Miller and Nicely was not identical to any of those used in the present study.
(2) Their filter cutoffs and slopes were slightly different.
(3) Their syllables were presented against a background of noise.
(4) Their syllables were presented by live voice whereas in this study the recorded syllables resulted in a type-token ratio of one.

In spite of these differences, there are very similar trends in the two sets of data, particularly with respect to the features nasal, voice, sibilant, and high/anterior.

The differential effects of high- and low-pass filtering on feature recognition become visible when performance functions are constructed for each feature. Although these functions do not take feature redundancy into account, they do show how the absolute level of performance on a feature changes across conditions. Feature recognition, expressed in terms of the percentage of information transmitted, is given in the Appendix for each syllable set. We will limit discussion here to some of the more interesting findings for the features identified most frequently by SINFA: sibilance, anterior, voice, and nasal.

Sibilance. Performance functions for sibilance in the CV-1 set are shown in Figure 2. High-pass filtering had very little effect on the recognition of this feature, while a steep drop in the function occurred for low-pass filtering. Very similar results were obtained for the other syllable sets. The high crossover point of the functions, 2800 Hz, indicates that cues for sibilant sounds lie in the high-frequency region of the spectrum, above 2000 Hz. This result is, of course, consistent with the findings of other studies (Hughes and Halle, 1956; Strevens, 1960).

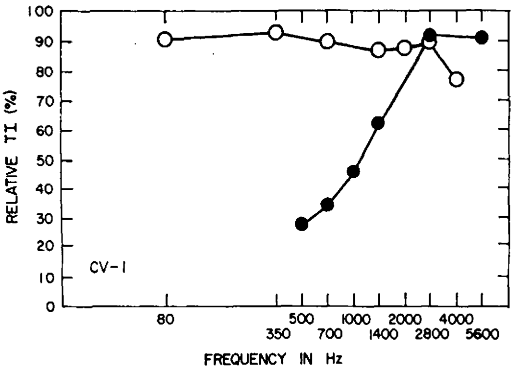

FIGURE 2. Percentage of information transmitted for the feature sibilance in the CV-1 syllable set as a function of filter cutoff. Filled symbols represent low-pass conditions. Open symbols represent high-pass conditions.

Anterior. Figure 3 shows performance for the feature anterior in the CV-2 and VC-2 sets, although anterior and high are indistinguishable in the VC-2 set. The function for high/anterior in the VC-1 set is very similar to that shown

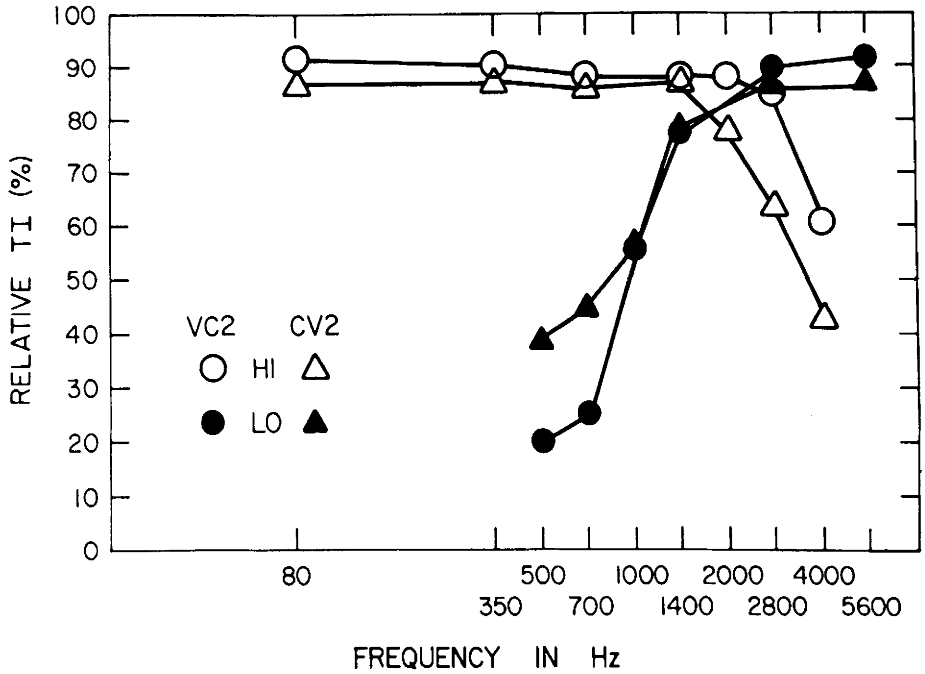

FIGURE 3. Percentage of information transmitted for the feature anterior in the CV-2 and VC-2 syllable sets as a function of filter cutoff.

for VC-2, while that for CV-1 is similar to the CV-2 function. For CV syllables, the crossover point, approximately 1700 Hz, is lower than that for VC syllables—about 2400 Hz. This suggests that the cues for high/anterior are partly dependent on the position of the consonant within the syllable.

Voice. Figure 4 shows performance for voice in the CV-2 and VC-2 sets. This feature was more adversely affected by high-pass than by low-pass filtering. Compared to other features, however, the drop in performance for voice, under the most severe high-pass conditions, was not very great. This result might have been predicted since there are a number of cues for voicing that are not strictly spectral. Among these are intensity (Strevens, 1960), duration of the consonant and vowel portions of the syllable (House and Fairbanks, 1953; Denes, 1955), and voice onset time (Lisker and Abramson, 1964).

Nasal. The nasality feature was distinctive only in the CV-2 and VC-2 sets, and its recognition functions are shown in Figure 5. Like voicing, nasality was more adversely affected by high-pass than by low-pass filtering. The crossover points were 850 Hz and 1150 Hz for the CV-2 and VC-2 sets respectively.

Comparison of Normal-Hearing and Hearing-Impaired Listeners

The effects of high- and low-pass filtering on consonant confusions, which are described above, bear a close resemblance to the effects of sensorineural

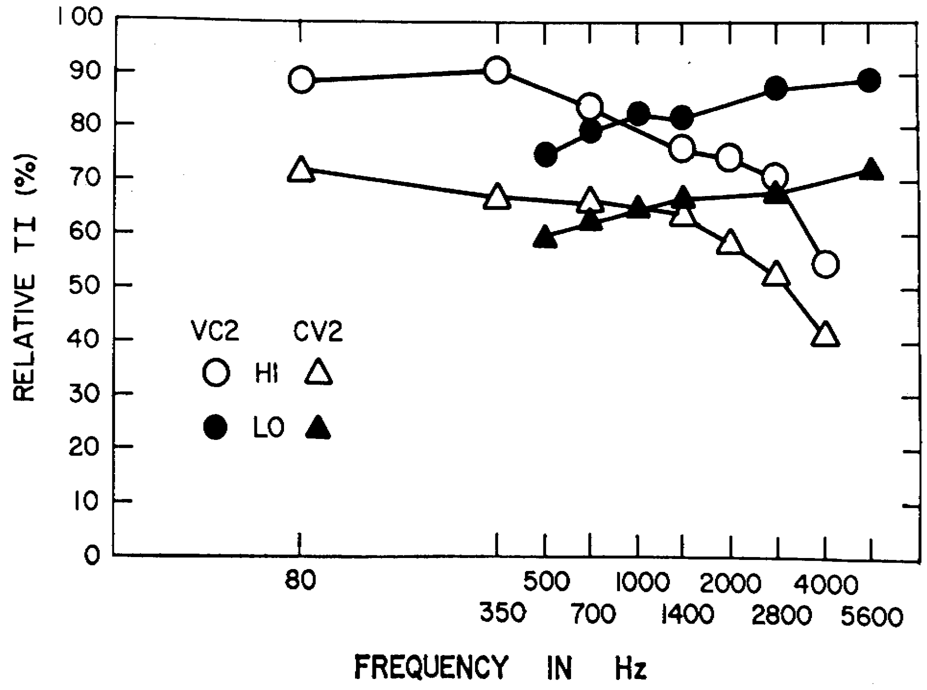

FIGURE 4. Percentage of information transmitted for the feature voice in the CV-2 and VC-2 syllable sets as a function of filter cutoff.

hearing loss on consonant confusions. As predicted, low-pass filtering produces patterns of consonant confusions comparable to those seen in patients with high-frequency hearing loss, and high-pass filtering produces patterns comparable to those seen in patients with moderate or severe flat hearing losses (see Bilger and Wang, 1976).

In that study, we defined a measure of intersubject similarity and used Johnson's (1973) pairwise multidimensional scaling procedure to represent the similarities spatially. By including the various filtering conditions as hypothetical "subjects" and scaling the resulting similarity matrix, we can show more clearly how the various conditions produce performance similar to that seen in the hearing-impaired listeners.

The similarity measure is obtained as follows: A large weight is assigned to the first feature identified by SINFA, and progressively smaller weights are assigned to features identified in successive iterations. The similarity measure is then obtained by summing the product of corresponding feature weights. For example, for the first column of Table 3, weights varying from six to one are assigned to the features sibilance through coronal, respectively. For the second column, weights from six to one are assigned to the features sibilance through place, respectively. The resulting similarity measure for these two conditions is 81. Although all features that two subjects have in common con-

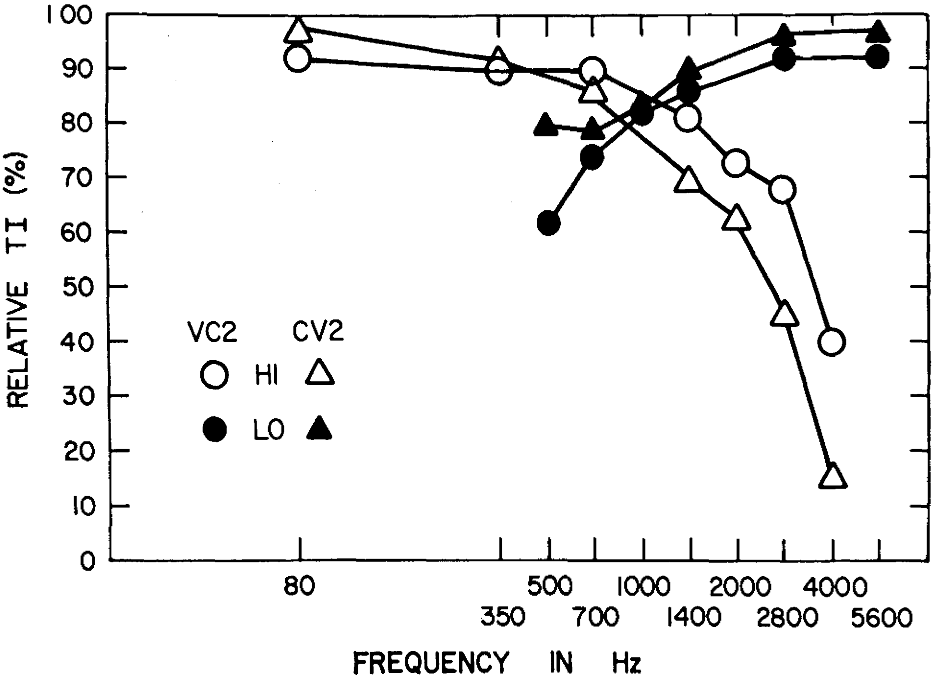

FIGURE 5. Percentage of information transmitted for the feature nasal in the CV-2 and VC-2 syllable sets as a function of filter cutoff.

tribute to the similarity measure, the greatest contribution is made by features identified in the first two or three iterations.

A similarity matrix was obtained in this way for each syllable set, for 23 hearing-impaired subjects and 12 hypothetical subjects corresponding to various filter conditions. In addition, a similarity matrix for all four sets combined was obtained by transforming the similarity measures to standard scores and summing across sets. The matrices were then scaled to produce five subject spaces. The configurations were rotated, if necessary, to increase their correspondence to the configurations reported earlier (Bilger and Wang, 1976).

The scaling solutions are shown in Figures 6-10. The values of θ, Johnson's lack-of-fit measure, indicate that a two-dimensional solution is adequate for the present purposes. In these figures, open symbols represent the hypothetical subjects (that is, the various filtering conditions), and closed symbols represent the hearing-impaired listeners. Closed circles represent listeners with normal hearing or mild hearing loss; triangles represent listeners with high-frequency hearing loss; and inverted triangles represent listeners with flat hearing loss. Identification numbers for individual subjects correspond to those used by Bilger and Wang (1976).

In Figure 6, which shows the configuration obtained for all syllable sets combined, the clearest differentiation is between the high-frequency-loss subjects

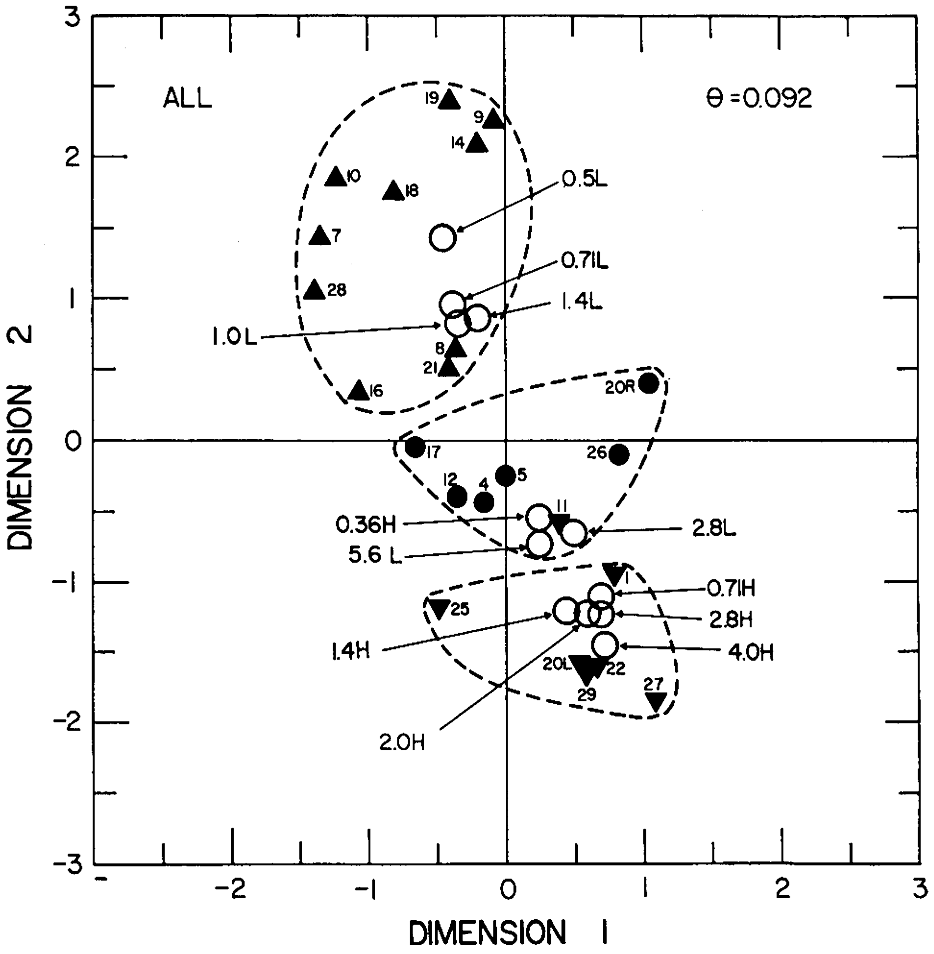

FIGURE 6. Scaling solution based on similarity of consonant confusions across all four syllable sets. In this and subsequent figures, closed symbols represent hearing-impaired listeners. Open symbols represent hypothetical subjects corresponding to different high-pass (H) and low-pass (L) conditions. Circles represent normal hearing or mild hearing loss, triangles represent high-frequency hearing loss, and inverted triangles represent flat hearing loss. θ, the lack-of-fit measure for the scaling solution, is shown in the upper right-hand corner.

and other listeners. In addition, the four most severe low-pass conditions fall clearly within this cluster. Although the remaining two clusters are not separated so clearly, there are only two misclassifications. The high-pass condition 710–5600 Hz falls within the flat-loss group, and one subject from that group resembles the mild filtering conditions more than expected.

Similar patterns are apparent in Figures 7–10, which show the configurations obtained for individual syllable sets. In each figure misclassifications of hearing-impaired listeners are apparent when a filled symbol appears in an inappropriate cluster. The cutoff at which filtering produces performance similar to that of a hearing-impaired group varies slightly from one syllable set to the next, but in no case does severe filtering produce performance similar to that of the inappropriate hearing-impaired group.

For the CV-1 set (see Figure 7), the two worst low-pass conditions fall within the high-frequency-loss group. At the other extreme, the worst high-pass condition falls near the flat-loss group. Although the remaining conditions have been included within the normal group, several clearly lie near the flat-loss group. This reflects the fact that sibilance and voice were well perceived under these conditions and were also well perceived by the normal and flat-loss groups.

For the VC-1 set (see Figure 8), the four worst low-pass conditions fall within the high-frequency-loss group, and the four worst high-pass conditions fall within the flat-loss group. For the mild filtering conditions, two fall within the

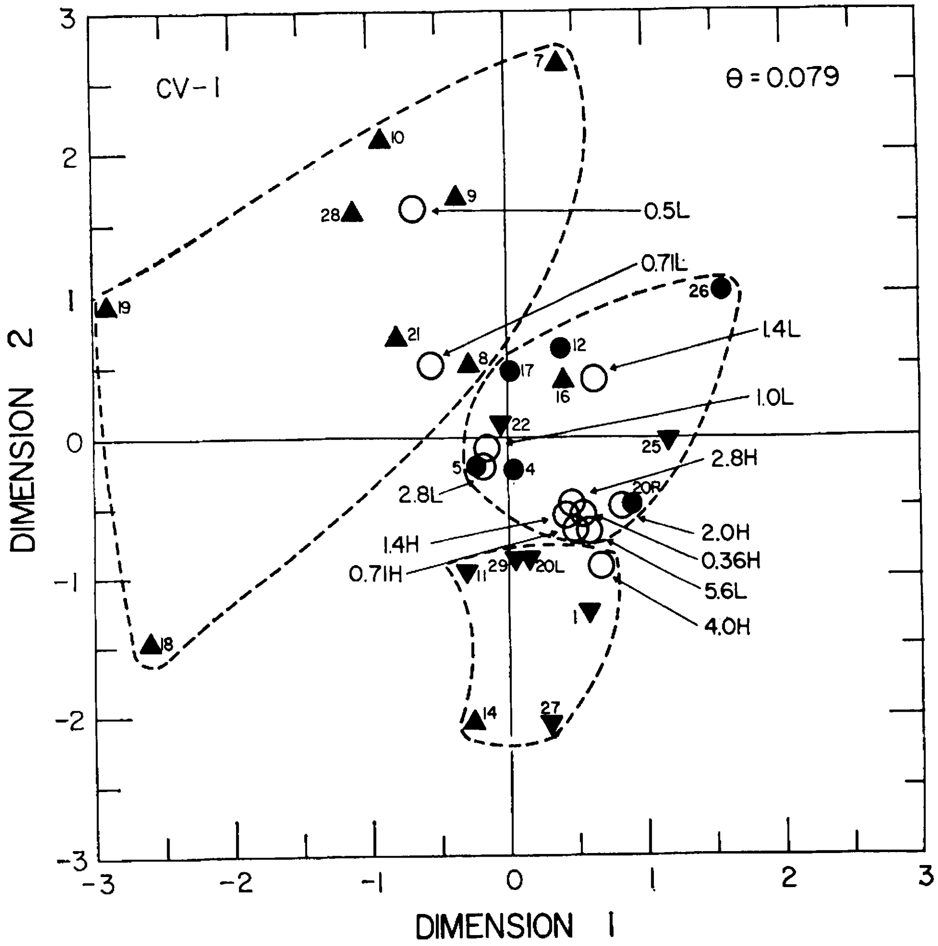

Figure 7. Scaling solution based on similarity of consonant confusions for the CV-1 syllable set.

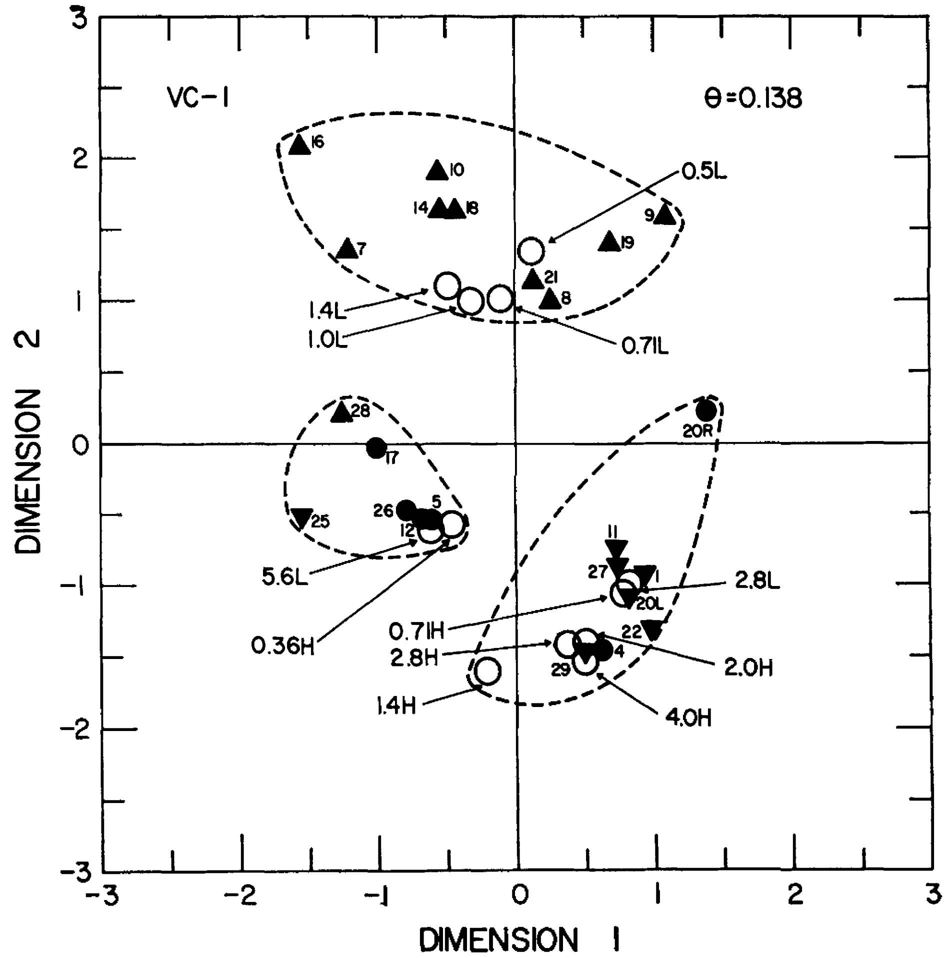

FIGURE 8. Scaling solution based on similarity of consonant confusions for the VC-1 syllable set.

normal group and two fall within the flat-loss group. The mild conditions that fall in the flat-loss group are the conditions in which duration, rather than frication, was identified in the second iteration by SINFA. There are also two listeners previously classified in the normal group for VC-1 who are now misclassified with the flat group.

Results for the CV-2 set, which are shown in Figure 9, contain many more misclassifications of listeners than results for other sets. Since this was also true when filter conditions were not included in the configuration (see Bilger and Wang, 1976), no attempt was made to improve the consistency of these clusters. Again, severe low-pass filtering produces patterns of consonant confusions similar to those seen in high-frequency-loss subjects, and high-pass filtering produces less systematic trends. It may be noted, however, that the two most severe high-pass conditions are those that appear in the low-right

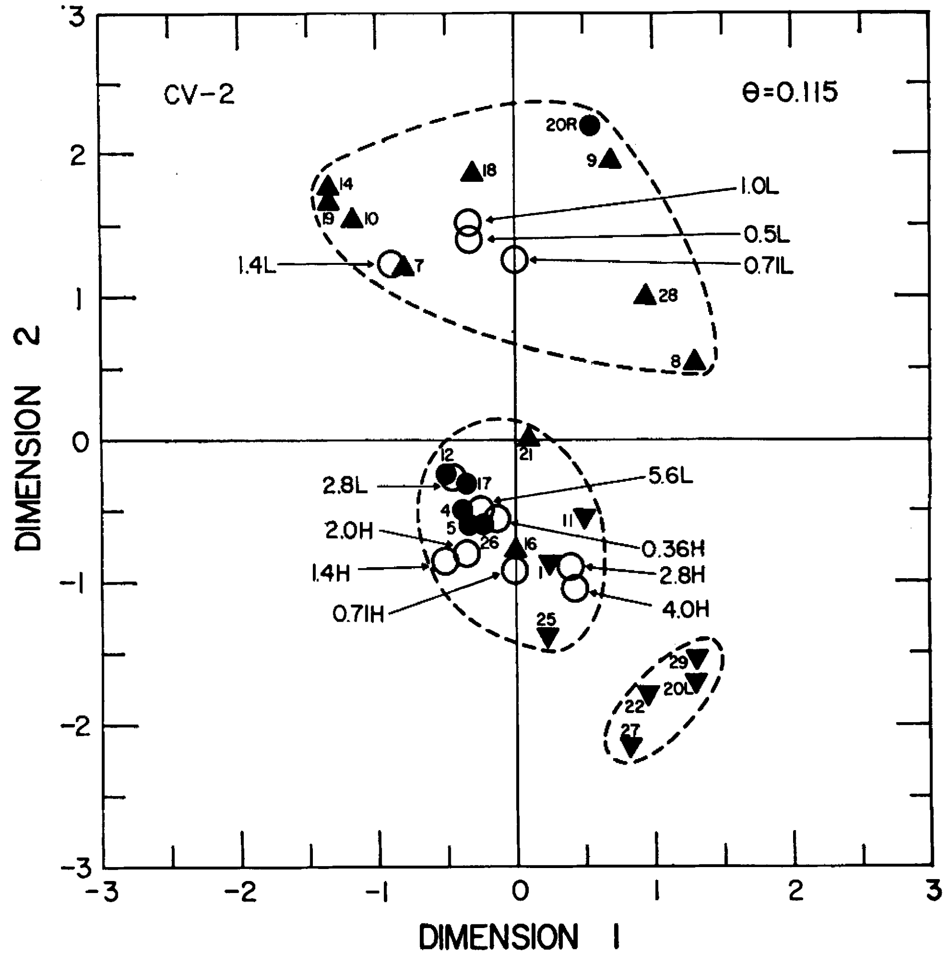

FIGURE 9. Scaling solution based on similarity of consonant confusions for the CV-2 syllable set.

quadrant of the figure, close to three of the misclassified flat-loss subjects.

For the VC-2 set (see Figure 10), low-pass filtering reproduces the result seen in the previous configuration. However, since the high-pass filtering produced very little change in the ranking of features (see Table 6) it is not surprising that these conditions fall close together in the configuration in Figure 10. Only the wide-band condition falls within the normal group.

To summarize, the effects of low-pass filtering at 1400 Hz or lower are to produce patterns of consonant confusions very similar to those seen in patients with high-frequency sensorineural hearing loss. The effects of high-pass filtering are less predictable; but, when the cutoff is raised to 2800 Hz, patterns of confusions resemble those of listeners with moderate or severe flat hearing losses.

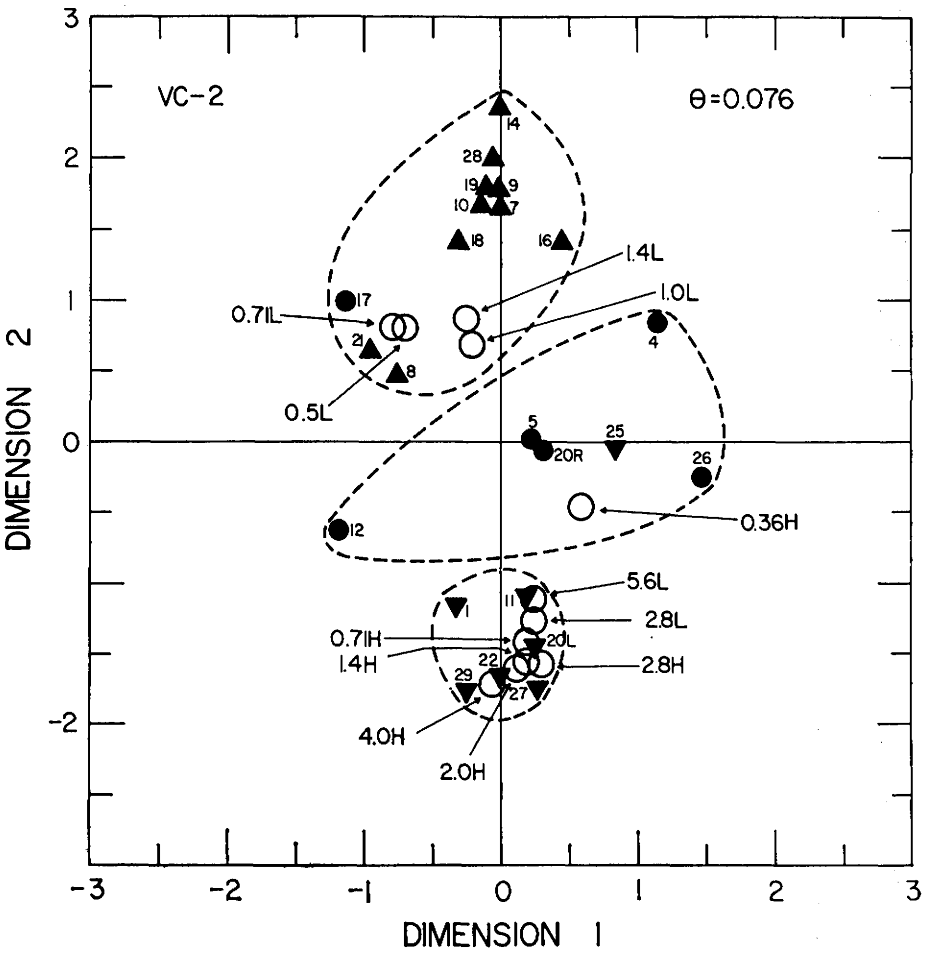

FIGURE 10. Scaling solution based on similarity of consonant confusions for the VC-2 syllable set.

Additional Comparisons of Normal-Hearing and Hearing-Impaired Listeners

Although the focus of this study is the comparison of hearing-impaired listeners and normal-hearing subjects listening to filtered speech, data for the four syllable sets are also available for normal-hearing subjects listening to undistorted speech and to speech masked by a broad-band noise (Wang and Bilger, 1973). Since it might be of interest to compare these groups with the hearing-impaired listeners, two additional analyses were performed. In the first, the consonant confusions of the hearing-impaired listeners were compared with those of six normal-hearing subjects listening to the syllables at 13 levels ranging from 20- to 115-dB SPL. In the second, the comparison was with four normal-hearing subjects listening to the syllables at six speech-to-noise ratios ranging from +15 to −10 dB. In each comparison, confusion matrices were obtained by pooling responses for all listeners for a given condition. Similarity matrices were generated and scaled as described previously. It should be noted, however, that the resulting configurations cannot be directly compared with those above or with each other because different stimulus conditions are included in different configurations.

The results for the first of these analyses are shown in Figure 11, where the data for normal-hearing listeners are shown as open circles. Each circle represents one of the 13 levels of signal presentation. The high-frequency-loss group is clearly separated from the others. The two conditions for normal-hearing subjects, which have been included arbitrarily in the cluster for flat-loss subjects, represent signal levels of 20- and 25-dB SPL. There is no tendency for

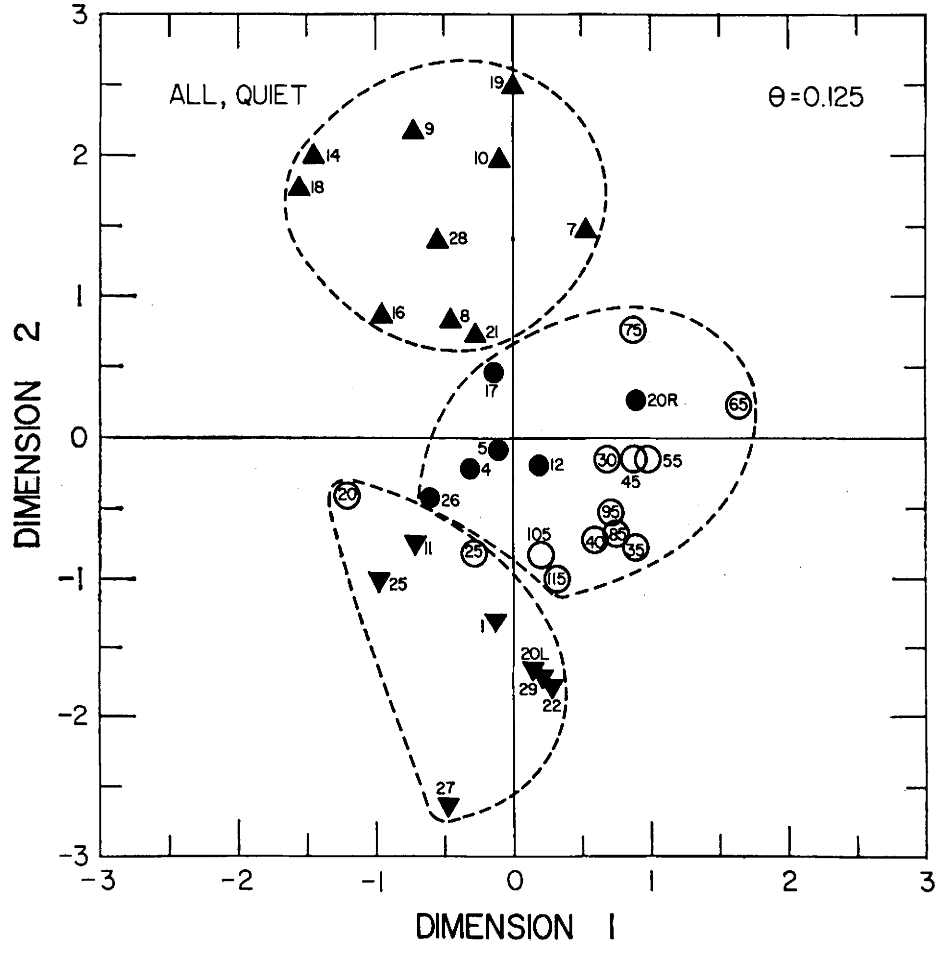

FIGURE 11. Scaling solution based on similarity of consonant confusions across all four syllable sets. Closed symbols represent hearing-impaired listeners; open symbols represent hypothetical subjects corresponding to listening conditions in which the level of speech varied from 20-dB to 115-dB SPL (Wang and Bilger, 1973).

high levels of presentation to produce abnormal patterns of consonant confusions. The analysis, therefore, provides a check on the possibility that some of the confusion patterns seen in hearing-impaired listeners might be caused by instrumental distortion of the stimulus on playback rather than distortion produced by the auditory lesion.

The comparison of noise conditions and hearing-impaired listeners is shown in Figure 12. The six noise conditions are represented as open circles. One condition, +15 dB speech-to-noise ratio, produces performance comparable to that of the listeners with relatively normal hearing. The remaining five conditions may be classified within the high-frequency-loss group, although they

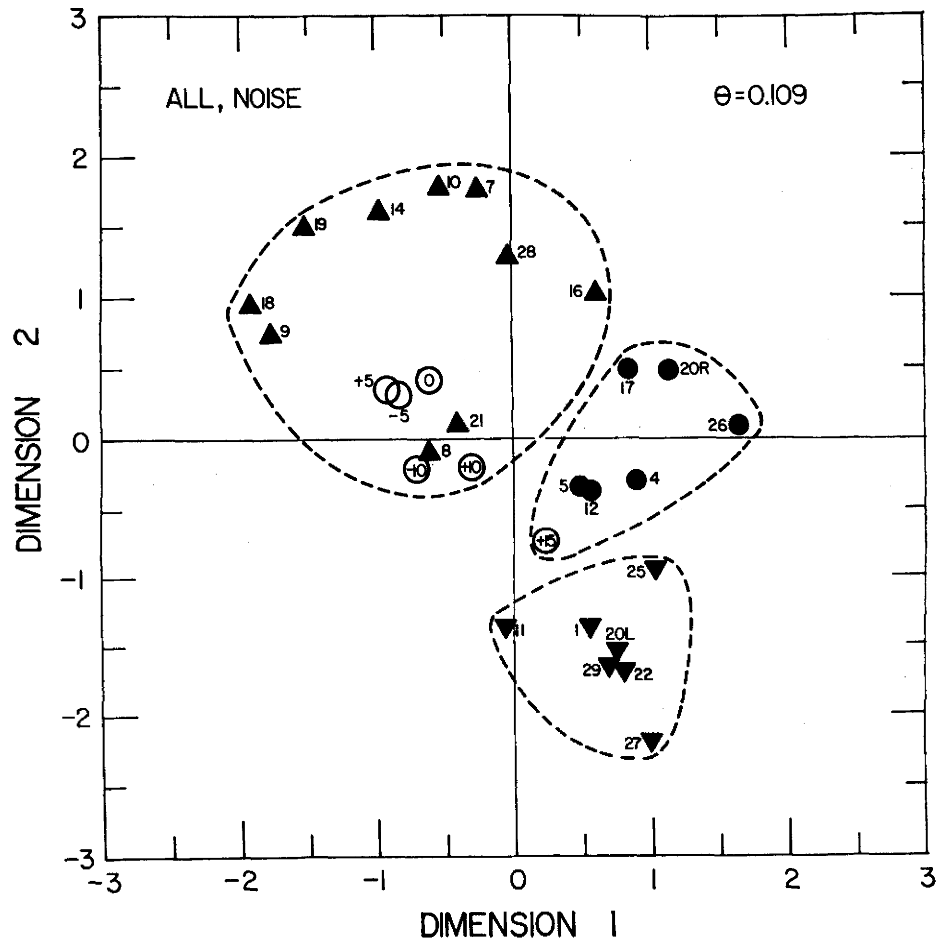

Figure 12. Scaling solution based on similarity of consonant confusions across all four syllable sets. Closed symbols represent hearing-impaired listeners; open symbols represent hypothetical subjects corresponding to listening conditions in which speech was presented with a broad-band masker and signal-to-noise ratios ranged from −10 dB to +15 dB (Wang and Bilger, 1973).

tend to form a subcluster that includes only two of the hearing-impaired listeners in this group. This is consistent with the finding of Miller and Nicely (1955) that broad-band masking noise gives results comparable to those obtained with low-pass filtering because of the effectiveness with which broad-band noise masks high-frequency information.

DISCUSSION

The results presented here indicate that the patterns of consonant confusions generated by subjects with sensorineural hearing loss are like those generated by normal-hearing subjects in response to the appropriate instrumental distortion of speech. Although there is some indication in Figure 6 that the filter conditions define a dimension along which the three groups of hearing-impaired listeners may be contrasted, we are reluctant to offer such an interpretation of Dimension 2 in Figures 6-10 for two reasons. First, in the configurations for individual syllable sets, the 12 filter conditions do not consistently fall along a single line. Thus the apparent "dimension" in Figure 6 may in part be an artifact of averaging similarities across different syllable sets. Second, and more importantly, we have used multidimensional scaling only because it provides a convenient way of representing similarities spatially, and not because we believe that it will reveal the underlying dimensions of a subject space. Thus we have chosen not to interpret dimensions of the resulting subject space, but rather to focus on the similarities in performance per se.

It should also be emphasized here that our results deal with the pattern of errors made by the subjects and not with the level of their overall performance (percentage of correct responses or of transmitted information). Previously we have reported (Bilger and Wang, 1976) that error rates and error patterns are relatively independent characteristics of performance on the consonant-recognition task used here. There we found that subjects with sensorineural hearing loss who demonstrated comparable patterns of consonant confusions can differ dramatically in their overall level of performance; and, conversely, that subjects with comparable levels of performance can generate drastically different patterns of consonant confusions. (Presumably, many previous attempts to show the similarity between high-frequency hearing loss and low-pass filtered speech have faltered because the traditional measure of speech perception, recognition of monosyllabic words, does not permit one to deal separately with the rate and pattern of errors.) Thus, although normal- and hard-of-hearing subjects show similar error patterns under appropriate conditions, overall levels of performance often differ between the groups. For a given filter condition, the normal-hearing subjects consistently make fewer recognition errors than hearing-impaired subjects whose audiograms approximate that filter condition. However, since we did not attempt to match audiograms and filter conditions closely, as did Sher and Owens (1974), this conclusion is at best a tentative one. On the other hand, the normal-hearing subjects in this study gave the same error patterns and rates for unfiltered speech (80–5600 Hz) that we obtained previously for normal-hearing control subjects (Bilger and Wang, 1976).

That the pattern of consonant confusions for subjects with sloping, high-frequency sensorineural hearing loss is like that for normal-hearing subjects listening to low-pass filtered speech, is clearly demonstrated in our results, and it seems intuitively obvious that this should be the case. That the pattern of confusions for subjects with flat sensorineural hearing loss is like that for normal-hearing subjects listening to high-pass filtered speech is demonstrated, albeit less clearly, and although the finding is explainable, it is not intuitively obvious why the subjects with flat losses should resemble high-pass filtering rather than unfiltered speech.

To explain this result, it is necessary to consider the spectrum of speech delivered to the subject in relation to his auditory sensitivity. This should allow us to specify the frequency spectrum of the speech heard by the subject in a particular situation. Figure 13 illustrates these points.

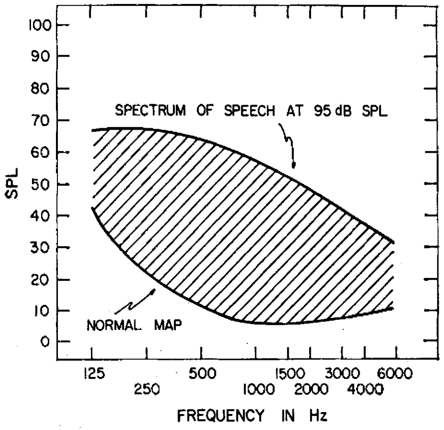

FIGURE 13. Comparison of average spectrum level of speech at 95-dB SPL and normal auditory threshold.

The normal sensitivity curve shown in Figure 13 was derived from current audiometric standards (ANSI, 1969) as adapted to the TDH-49 earphone. The average spectrum levels for speech were taken from data by Benson and Hirsh (1953). The spectrum levels plotted in Figure 13 correspond to the 95-dB overall level of speech that was used in the present experiments. The shaded area between these two curves is a representation of the speech signals heard by the subject with normal hearing when the speech signal is not filtered. If the critical ratio for each frequency had been added to the spectrum level of the speech, then the figure would be a first approximation to the Articulation Index (French and Steinberg, 1947).

The effect of filtering the speech signal presented to a normal-hearing subject can be visualized quite easily in terms of Figure 13. For example, the effect of low-pass filtering is to remove high-frequency cues to speech, just as the effect of high-frequency hearing loss is to remove comparable high-frequency information about the speech signal. In the former case the spectrum of speech changes; in the latter case the sensitivity curve is altered. High-pass filtering of speech presented to normal-hearing subjects has an equally obvious effect, one of removing low-frequency information from the speech signal. Because the effect of a flat sensorineural hearing loss is to displace the entire sensitivity curve upward, it is more difficult to conceptualize its exact effect on the spectrum level of speech heard by the subject. If the data are replotted in terms of the distance, in decibels, between the sensitivity curve and the speech spectrum level curve, with a correction for the critical ratio (Reed and Bilger, 1973), then the distortion that arises from the fact that both functions are curvilinear can be removed. The resultant decibel number is essentially a sensation level.

Curves showing the average sensation level of the spectrum level of speech are shown in Figure 14 for the three groups of subjects reported by Bilger and Wang. To obtain these curves, the sensation level of the speech was calculated for each of the subjects at each audiometric frequency by subtracting the pure-tone threshold from the spectrum level of speech for the overall level at which it had been presented and then adding the critical ratio for that fre-

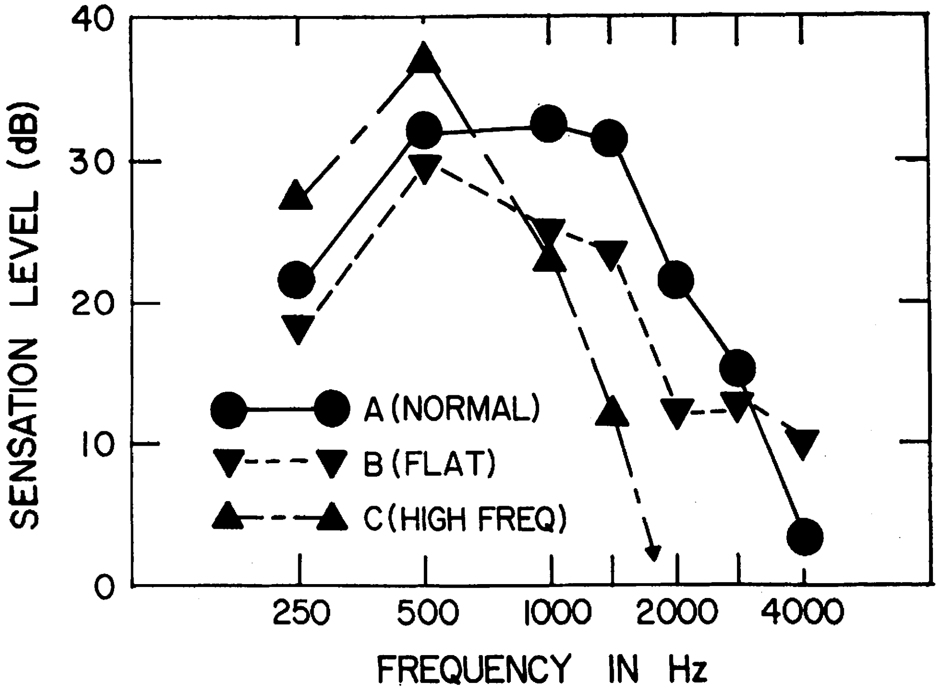

FIGURE 14. Sensation level of speech at selected frequencies for three groups of listeners studied by Bilger and Wang (1976).

quency. The resultant sensation levels were then averaged across subjects in each group. The function relating sensation level of the speech to audiometric frequency for the six normal-hearing subjects (circles) is the reference against which to evaluate the functions for the other two groups of subjects. The function for the 10 subjects with high-frequency losses (triangles) lies above that for normal-hearing subjects at 250 and 500 Hz, but drops precipitously above 500 Hz. This relation between the functions is consistent with the subjects' results on the consonant recognition task and with the results seen for low-pass filtering, in that nasality and voice were well perceived but sibilance was poorly perceived in these situations.

The contrast between the normal function and that for seven subjects with flat or rising audiograms (inverted triangles in Figure 14) is less dramatic. Between 250 and 2000 Hz, the function for the flat-loss group is essentially parallel to the function for normal-hearing subjects and lies an average of 6 dB below it; from 2800 to 5600 Hz the functions for the two groups are essentially the same. (This latter similarity arises because several of these subjects had better hearing at 2000, 4000, or 8000 Hz than at lower frequencies.) To understand why this attenuation of the frequencies below 2000 Hz effectively acts like a high-pass filter, one needs to refer to the work reported by French and Steinberg (1947), especially their Figure 3 and Table 1. Their Figure 3 shows that the distribution of spectrum levels that occur in speech is much narrower for the frequency region from 250 to 500 Hz than it is in the higher frequency regions. In their Table 1, they also show the difference between the long-term average spectrum levels of speech and the peak levels that are reached in only 1% of 0.125-sec samples. Again, these data show that this difference is smaller in the frequency region below 500 Hz than it is at higher frequencies. Thus, a small decibel deficit at 250 or 500 Hz is probably much more detrimental to the perception of low-frequency speech cues than a corresponding deficit at 2000 Hz is to the perception of high-frequency speech cues.

The consonant confusions made by our flat-loss group and by our normal-hearing subjects listening to high-pass filtered speech can be characterized by the relatively poor perception of nasality and the excellent perception of sibilance. Thus it seems that nasality, and possibly voicing, are the specific low-frequency features that were most adversely affected either by high-pass filtering or a flat sensorineural hearing loss. In the case of the subjects with flat sensorineural hearing loss, however, the effect is not a simple filtering effect, but rather an effect which involves the interaction between the frequency dependence of the level distributions in speech and the rule of setting the overall level of the speech signal 40 dB above the speech reception threshold. The similarity between subjects with flat sensorineural hearing losses and normal subjects listening to high-pass filtered speech may well be an artifact of the 40-dB rule.

Our results suggest that a rule for setting the overall level of speech in a recognition task (such as speech-discrimination testing) that ignores the listener's audiometric configuration introduces certain systematic biases into the results. The nature of the biases is such that the rule leads to the presentation of the speech signals at too high a level (especially regarding spectrum levels at 250 and 500 Hz) for listeners with sloping, high-frequency losses (by about 5 dB), and at too low a level (especially at 250 and 500 Hz) for listeners with flat sensorineural losses (by about 6 dB). The existence of these biases could have been predicted from the earlier results of Fletcher (1950) and of French and Steinberg (1947), which show that the frequency range necessary for responding to spondaic words is different from that necessary to identify monosyllabic words or consonants.

ACKNOWLEDGMENT

This research was supported by grants from NINCDS to the University of Pittsburgh (NS 4105) and to the Eye and Ear Hospital of Pittsburgh (NS 12501). In addition, much of the data analysis was also supported by faculty research funds made available to M. D. Wang through the Social Sciences Statistics Center of the University of Maryland Baltimore County. Many of the facilities used in this research are the result of gifts from the Richard King Mellon Charitable Trusts, Inc., and from the Edwards-Lazear Foundation of Pittsburgh. We wish to acknowledge the invaluable contributions of our colleagues and friends to this work: Richard Wolf and Joseph Wisniewski for their help in setting up the equipment and keeping it operational; Linet Feigel and Walt Jesteadt for their assistance in programming and data analysis; Louis Braida for his helpful comments on an earlier version of this manuscript. Requests for reprints should be addressed to Marilyn D. Wang, Department of Psychology, UMBC, 5401 Wilkens Avenue, Baltimore, Maryland 21228; Charlotte M. Reed, Massachusetts Institute of Technology, Room 36-747, Cambridge, Massachusetts 02139; Robert C. Bilger, Department of Speech and Hearing Sciences, University of Illinois, 901 South Sixth Street, Champaign, Illinois 61820.

REFERENCES

American National Standards Institute, *Specification for Audiometers.* S3.6-1969. New York: American National Standards Institute (1970).

Benson, R. W., and Hirsh, I. J., Some variables in audio spectrometry. *J. acoust. Soc. Am.,* **25**, 499-505 (1953).

Bilger, R. C., and Wang, M. D., Consonant confusions in patients with sensorineural hearing loss. *J. Speech Hearing Res.,* **19**, 718-748 (1976).

Denes, P., Effect of duration on the percepton of voicing. *J. acoust. Soc. Am.,* **27**, 761-767 (1955).

Fletcher, H., A method of calculating hearing loss for speech from an audiogram. *J. acoust. Soc. Am.,* **22**, 1-5 (1950).

French, N. R., and Steinberg, J. C., Factors governing the intelligibility of speech sounds. *J. acoust. Soc. Am.,* **19**, 90-119 (1947).

Hirsh, I. J., Reynolds, E. G., and Joseph, M., Intelligibility of difficult speech materials. *J. acoust. Soc. Am.,* **26**, 530-538 (1954).

House, A. S., and Fairbanks, G., The influence of consonant environment upon the secondary acoustical characteristics of vowels. *J. acoust. Soc. Am.,* **25**, 105-115 (1953).

Hughes, G., and Halle, M., Spectral properties of fricative consonants. *J. acoust. Soc. Am.,* **28**, 303-310 (1956).

Johnson, R. M., Pairwise nonmetric multidimensional scaling. *Psychometrika,* **38**, 11-18 (1973).

LISKER, L., and ABRAMSON, A. S., A cross-language study of voicing in initial stops: Acoustical measurements. *Word,* **20**, 384-422 (1964).

MILLER, G. A., and NICELY, P. E., Analysis of perceptual confusions among some English consonants. *J. acoust. Soc. Am.,* **27**, 338-352 (1955).

OWENS, E., BENEDICT, M., and SCHUBERT, E. D., Consonant phonemic errors associated with pure-tone configurations and certain kinds of hearing impairments. *J. Speech Hearing Res.,* **15**, 308-322 (1972).

REED, C. M., and BILGER, R. C., A comparative study of S/No and E/No. *J. acoust. Soc. Am.,* **53**, 1039-1044 (1973).

SHER, A. E., and OWENS, E., Consonant confusions associated with hearing loss above 2000 Hz. *J. Speech Hearing Res.,* **17**, 669-681 (1974).

STREVENS, P., Spectra of fricative noise in human speech. *Lang. Speech,* **3**, 32-49 (1960).

WANG, M. D., and BILGER, R. C., Consonant confusions in noise: A study of perceptual features. *J. acoust. Soc. Am.,* **54**, 1248-1266 (1973).

Received February 28, 1977.
Accepted August 2, 1977.

APPENDIX

Percentage of feature information transmitted for each filter condition in each syllable set.

CV-1 Set	*Filter Condition (Hz)*					
Feature	*80-5600*	*80-2800*	*80-1400*	*80-1000*	*80-710*	*80-500*
Voice	0.897	0.918	0.885	0.806	0.735	0.584
Frication	0.635	0.657	0.488	0.370	0.325	0.287
Place	0.694	0.702	0.513	0.288	0.168	0.133
Duration	0.771	0.758	0.511	0.377	0.244	0.205
High/anterior	0.885	0.885	0.685	0.374	0.202	0.163
Back	0.751	0.763	0.507	0.225	0.188	0.139
Coronal	0.505	0.531	0.332	0.185	0.115	0.108
Continuant	0.587	0.589	0.449	0.311	0.243	0.221
Strident	0.453	0.476	0.365	0.277	0.218	0.194
Sibilant	0.914	0.932	0.624	0.458	0.340	0.278
	355-5600	*710-5600*	*1400-5600*	*2000-5600*	*2800-5600*	*4000-5600*
Voice	0.886	0.844	0.803	0.724	0.710	0.596
Frication	0.631	0.628	0.571	0.566	0.574	0.467
Place	0.693	0.685	0.664	0.629	0.569	0.296
Duration	0.790	0.777	0.782	0.771	0.710	0.568
High/anterior	0.875	0.853	0.823	0.761	0.706	0.337
Back	0.732	0.679	0.613	0.551	0.517	0.141
Coronal	0.517	0.510	0.494	0.501	0.465	0.311
Continuant	0.586	0.594	0.555	0.535	0.499	0.375
Strident	0.479	0.495	0.443	0.463	0.512	0.417
Sibilant	0.926	0.895	0.872	0.885	0.904	0.767

VC-1 Set	*Filter Condition (Hz)*					
Feature	*80-5600*	*80-2800*	*80-1400*	*80-1000*	*80-710*	*80-500*
Voice	0.842	0.856	0.831	0.818	0.796	0.784
Frication	0.925	0.938	0.865	0.770	0.572	0.470
Place	0.831	0.828	0.671	0.412	0.193	0.135
Duration	0.919	0.930	0.683	0.482	0.275	0.228
High/anterior	0.891	0.895	0.781	0.477	0.216	0.173
Back	0.865	0.869	0.748	0.523	0.446	0.304
Coronal	0.805	0.798	0.601	0.403	0.225	0.143
Continuant	0.893	0.917	0.844	0.747	0.583	0.467
Strident	0.787	0.789	0.618	0.542	0.347	0.296
Sibilant	0.956	0.962	0.733	0.534	0.296	0.270
	355-5600	*710-5600*	*1400-5600*	*2000-5600*	*2800-5600*	*4000-5600*
Voice	0.890	0.817	0.683	0.708	0.635	0.527
Frication	0.916	0.926	0.848	0.848	0.746	0.579
Place	0.845	0.814	0.799	0.823	0.782	0.539
Duration	0.937	0.949	0.929	0.922	0.921	0.841
High/anterior	0.894	0.875	0.890	0.932	0.899	0.655
Back	0.848	0.835	0.843	0.906	0.865	0.471
Coronal	0.827	0.792	0.744	0.746	0.697	0.466
Continuant	0.893	0.915	0.846	0.829	0.746	0.565
Strident	0.810	0.771	0.712	0.697	0.637	0.452
Sibilant	0.967	0.972	0.947	0.965	0.946	0.914

CV-2 Set	*Filter Condition (Hz)*					
Feature	*80-5600*	*80-2800*	*80-1400*	*80-1000*	*80-710*	*80-500*
Vocalic	0.956	0.966	0.897	0.701	0.526	0.506
Consonantal	0.854	0.842	0.810	0.623	0.481	0.346
High	0.724	0.681	0.658	0.461	0.296	0.293
Back	0.432	0.346	0.362	0.287	0.190	0.139
Anterior	0.868	0.854	0.787	0.558	0.450	0.387
Coronal	0.942	0.925	0.683	0.439	0.291	0.209
Continuant	0.685	0.693	0.643	0.514	0.397	0.402
Nasal	0.968	0.958	0.901	0.826	0.794	0.807
Strident	0.717	0.733	0.690	0.559	0.456	0.373
Voice	0.721	0.672	0.674	0.653	0.626	0.608
Round	0.862	0.818	0.789	0.550	0.335	0.186
Low	0.547	0.474	0.487	0.423	0.427	0.313
Place	0.864	0.836	0.697	0.485	0.369	0.284
Duration	0.976	0.961	0.648	0.488	0.325	0.230
Sibilant	0.980	0.965	0.739	0.606	0.458	0.359

	355-5600	710-5600	1400-5600	2000-5600	2800-5600	4000-5600
Vocalic	0.945	0.916	0.865	0.757	0.654	0.273
Consonantal	0.833	0.853	0.875	0.785	0.605	0.387
High	0.675	0.658	0.640	0.607	0.566	0.384
Back	0.318	0.304	0.294	0.254	0.146	0.056
Anterior	0.867	0.864	0.866	0.772	0.643	0.429
Coronal	0.913	0.917	0.851	0.779	0.693	0.421
Continuant	0.682	0.710	0.634	0.627	0.574	0.357
Nasal	0.918	0.858	0.700	0.634	0.448	0.146
Strident	0.735	0.731	0.679	0.664	0.628	0.459
Voice	0.674	0.655	0.651	0.590	0.539	0.411
Round	0.813	0.870	0.880	0.763	0.420	0.162
Low	0.467	0.472	0.466	0.441	0.389	0.267
Place	0.839	0.841	0.802	0.753	0.690	0.428
Duration	0.979	0.985	0.977	0.967	0.970	0.814
Sibilant	0.983	0.977	0.979	0.961	0.945	0.838

VC-2 Set			*Filter Condition (Hz)*			
Feature	*80-5600*	*80-2600*	*80-1400*	*80-1000*	*80-710*	*80-500*
Voice	0.887	0.868	0.823	0.829	0.802	0.751
Frication	0.925	0.910	0.854	0.766	0.627	0.453
Place	0.845	0.816	0.648	0.461	0.222	0.163
Duration	0.922	0.931	0.656	0.476	0.258	0.195
Nasal	0.925	0.916	0.856	0.825	0.744	0.619
High/anterior	0.925	0.898	0.781	0.559	0.246	0.195
Back	0.889	0.871	0.705	0.566	0.500	0.341
Coronal	0.796	0.761	0.547	0.398	0.229	0.148
Continuant	0.894	0.896	0.835	0.742	0.619	0.454
Strident	0.809	0.768	0.611	0.535	0.372	0.292
Sibilant	0.966	0.950	0.698	0.535	0.300	0.238

	355-5600	*710-5600*	*1400-5600*	*2000-5600*	*2800-5600*	*4000-5600*
Voice	0.907	0.836	0.761	0.752	0.708	0.552
Frication	0.916	0.854	0.797	0.718	0.662	0.426
Place	0.816	0.799	0.774	0.766	0.707	0.485
Duration	0.910	0.918	0.946	0.909	0.912	0.858
Nasal	0.902	0.899	0.813	0.728	0.676	0.397
High/anterior	0.898	0.881	0.880	0.875	0.848	0.607
Back	0.868	0.850	0.823	0.824	0.797	0.391
Coronal	0.768	0.742	0.709	0.698	0.608	0.416
Continuant	0.883	0.842	0.799	0.719	0.673	0.456
Strident	0.811	0.735	0.701	0.660	0.582	0.376
Sibilant	0.960	0.951	0.973	0.952	0.939	0.908