CONSONANT CONFUSIONS IN PATIENTS WITH SENSORINEURAL HEARING LOSS

ROBERT C. BILGER

University of Pittsburgh, Pennsylvania

MARILYN D. WANG

University of Maryland Baltimore County

Consonant confusion matrices were obtained from 22 outpatient listeners with sensorineural hearing loss for four sets of CV and VC nonsense syllables, presented monaurally at SRT + 40 dB. Testing was typically conducted for six hours on each of two separate days. Overall performance and patterns of confusions were stable over time. Analysis of the matrices in terms of phonological features indicated that the patterns of consonant confusions varied both with degree and configuration of the subject's loss. Scaling of intersubject similarity using a pairwise multidimensional scaling analysis resulted in consistent classification of subjects according to audiometric configuration into three groups—essentially normal hearing, flat or rising audiograms, and high-frequency hearing losses.

Although it is generally acknowledged that sensorineural hearing impairment is often accompanied by a loss of speech-recognition ability (typically measured as the percentage of monosyllabic words correctly recognized), relatively little research has been carried out documenting the nature of the speech-recognition loss. Most clinical research on speech perception has been concerned with the prediction of word-recognition scores and speech reception thresholds from audiometric data; the comparability of various clinical tests of word-recognition with different types of patients and different testing conditions; or the effects of instrumental distortions, such as noise or filtering, on word-recognition performance. A common finding of studies in the first category has been that word-recognition performance is not well predicted by audiometric data. However, the dependent variable analyzed has almost always been the patient's level of performance, that is, a word-recognition score. While such scores have utility for estimating the degree of handicap a patient suffers, they do little to illuminate the nature of the speech-recognition problem.

For both practical and theoretical reasons, it is important to examine the kinds of errors a patient makes (when he makes errors) and to determine whether audiometric configuration, or other variables, are related to the errors in a systematic way. Only recently has there begun to appear research describing phonemic confusions in patients with hearing loss and relating the confusions to other characteristics of the patient.

Oyer and Doudna (1959) analyzed errors made by patients with either conductive or nonconductive losses while responding to W-22 lists. They found that the two groups showed similar patterns of confusion, although the conductive patients were more consistent in their errors over time. They also noted that sound omissions and insertions were more frequent in word-final position than in word-initial position. In a similar study contrasting patients with acoustic trauma, Meniere's syndrome, presbycusis, and sensorineural loss, Schultz (1964) found that phonemic confusions were so infrequent and so idiosyncratic, both within and between diagnostic groups, that their usefulness for diagnostic purposes was not supported. Lawrence and Byers (1969) also reported idiosyncratic confusions for individual patients. They gathered extensive confusion data on five subjects with high-frequency hearing losses. The listeners' task was to identify consonant-vowel nonsense syllables formed by combining the voiceless fricatives /ʃ, s, f, θ/ with vowels /i, e, o, u/. They did note, however, that confusion patterns for the individual listeners stabilized rather quickly, with the largest number of errors occurring in the first testing session.

Extensive analyses of phonemic errors have been made by Owens and his colleagues (Owens and Schubert, 1968; Owens, Benedict, and Schubert, 1972; Sher and Owens, 1974). Using a multiple-choice word-recognition test, they compared phonemic error rates and phonemic confusions for patients with distinctly different audiometric configurations. Owens et al. (1972) found that identification of /s/ in both initial and final position, and /t/ and /θ/ in initial position seemed highly dependent on the frequency range above 2000 Hz, and that identification of initial and final /ʃ, tʃ, dʒ/ was heavily dependent on the frequency range between 1000 and 2000 Hz. They also noted that across all configurations the most frequently occurring confusions were the same. Thus, the likelihood that an error would occur was dependent, at least for some phonemes, on the configuration of the audiogram, but the specific error which was most likely to occur was the same for all groups. Of especial interest was the finding that the performance of normal-hearing subjects listening through a 780-Hz low-pass filter was highly similar to that of the comparable patient group. Subsequently, Sher and Owens (1974) confirmed that normal-hearing subjects listening to speech low-pass filtered at 2000 Hz could not be distinguished from listeners with a comparable high-frequency hearing loss with respect to phonemic error rates or phonemic confusions.

The research reported here was designed to explore systematically the nature of consonant confusions in patients with sensorineural hearing loss. All of the consonant phonemes of English were studied using both CV and VC nonsense syllables as stimuli, and employing a 16-alternative forced-choice response task. Nonsense syllables were used, rather than words, in order to maximize the contribution of acoustical factors to confusions and to minimize the contribution of linguistic factors.

Specifically, the study was concerned with three questions. First, are confusions sufficiently stable, over time and stimulus sets, that it is possible to describe, in a general way, the nature of a single patient's discrimination loss? To answer this, we obtained consonant confusions for four different sets of nonsense syllables on two occasions. Consistency of performance over time and stimulus sets was evaluated by a comparison of phonemic-error rates and phonemic confusions. Secondly, to what extent are confusion patterns idiosyncratic, and to what extent do patients fall into natural groups on the basis of these patterns? Our approach to this question was to compare the feature analyses of different patients with one another and to derive a measure of intersubject similarity based on these analyses. An advantage of feature analysis in this context is that it incorporates both the information about error rates for specific phonemes and the information about specific confusions into a single analysis and it permits asymmetries in the confusion matrix to be taken into account. Multidimensional scaling of the similarity measures was then used to determine whether patients showed a tendency to group themselves in a systematic way. If we could show that patients do tend to group themselves, a third question we wished to explore was whether such groups might not have other characteristics in common such as audiometric configuration.

METHOD

Subjects

The subjects were 22 patients, recruited from among those seen in the Audiology Department of Eye and Ear Hospital of Pittsburgh. We attempted to obtain subjects with elevated SRTs, reduced W-22 scores, and a variety of pure-tone audiometric configurations. A summary of descriptive and audiometric data for each subject is presented in Table 1. Three subjects (04, 05, and 20R) showed no evidence of hearing impairment and were included as normal control subjects. Only one subject (20) was tested in both ears.

Syllable Sets

Four different syllable sets were used. Each consisted of 48 nonsense syllables formed by combining 16 consonants with three vowels /i, a, u/. The consonants included in each set are shown in Table 2. Across all four sets a total of 129 different syllables was tested. These syllables represent all of the phonologically permissible CV and VC combinations of English consonants with the three vowels used.

Speech System

The speech system used for presentation of the syllables has been described

TABLE 1. Descriptive and audiometric data for subjects.

Subject	Age	Sex	Retest Interval	Pure-Tone Thresholds (ANSI, 1969) 250 Hz	500 Hz	1 kHz	2 kHz	4 kHz	8 kHz	Speech SRT	W-22
01	55	F	6 days	50 dB	45 dB	50 dB	50 dB	25 dB	15 dB	26 dB	92%
04	35	F	4 mos.	10	5	0	5	10	15	−2	100
05	34	F	15 mos.	5	5	15	5	5	0	−2	100
07	49	F	4 mos.	10	5	0	10	50	40	0	100
08	59	F	6 mos.	65	70	75	75	95	90	65	88
09	53	M	7 mos.	25	40	55	85	110	90	45	70
10	47	F	4 mos.	5	5	10	15	60	45	2	96
11	54	F	5 days	45	60	65	85	90	90	60	80
12	42	F	1 mo.	15	25	20	20	30	15	16	96
14	56	M	–	15	15	15	55	85	90	16	92
16	57	M	4.5 mos.	5	0	10	15	35	50	8	100
17	54	F	1 mo.	20	25	30	40	60	60	28	88
18	47	M	13 days	0	15	95	110	110	110	24	50
19	55	M	1 day	10	15	5	45	70	35	2	84
20R	63	F	–	15	5	0	0	20	10	0	100
20L	63	F	2 mos.	65	60	55	50	65	60	50	38
21	49	F	–	50	40	45	55	100	90	40	54
22	47	M	5 wks.	55	50	55	60	65	45	52	58
25	57	M	7.5 mos.	25	15	15	60	65	60	15	74
26	53	M	6.5 mos.	35	30	20	35	30	50	22	82
27	21	F	2-4 wks.	75	65	65	40	30	65	48	–
28	46	M	6 mos.	15	25	15	60	65	50	20	88
29	52	M	6.5 mos.	55	55	55	50	20	35	48	88

TABLE 2. 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/

in detail by Wang and Bilger (1973). Briefly, each of the 129 nonsense syllables was permanently recorded, by a male speaker, on the addressable vocabulary drum of a Cognitronics Speechmaker. A specially designed control system operated the Speechmaker in accordance with specifications provided by the experimenter for each trial. Output from the Speechmaker was calibrated by matching the average peak VU reading obtained for the syllables with that of a 1000-Hz tone. The speech signal was gated by an electronic switch (GS-1287) with a rise time of 1 msec, amplified, attenuated, and fed to a passive four-way splitter. (Although subjects were almost always tested individually, the system was capable of testing four subjects simultaneously.) The signal was then delivered to a console housed in a sound-treated booth (IAC-401A), from which it was led to a single TDH-49 earphone mounted in an MX-41/AR cushion. An identical phone was used as a dummy for the contralateral ear. Two subjects (20L and 27), however, required contralateral masking in the better ear. For these subjects, a broad-band masker (GS-455C) was amplified, attenuated, and delivered continuously through the dummy phone.

Procedure

A single test run consisted of 96 items from one syllable set. Syllables were presented at approximately 40 dB above the subject's SRT, unless the maximum of the system (120 dB) was reached or the subject requested a lower level. Within each run, each of the 48 different syllables in a set (16 consonants × 3 vowels) was presented twice, and the order of the syllables was completely random.

The subject was seated before a response console with a 4 × 4 array of response buttons, each labeled with a different consonant sound in conventional orthography. To the left of this array, a list of monosyllabic cue words was available for reference throughout the test session. Before each new syllable set was introduced, the experimenter illustrated the sound corresponding to each response button, referring the subject to the cue words when necessary. In addition to the response buttons, there were three coincidence indicators labeled *Warning, Observe*, and *Answer*. On each trial of a test run, there was a 500-msec warning interval, and a 511-msec observation interval during which the test syllable was presented. The subject responded by pressing one of the 16 response buttons. A 200-msec feedback interval followed, during which a green light in the upper portion of the correct response button was lit.

Each test day was divided into a morning and an afternoon session separated by a one-hour lunch break. In the morning the subject completed three test runs on each syllable set in turn. The order of the sets was varied so that CV and VC sets always alternated. The order of presentation of the sets was varied systematically across subjects, although strict counterbalancing was not achieved. In the afternoon, two additional test runs were completed for each syllable set.

Each subject was invited to return for a second day of testing. Retest intervals varied from one day to 15 months (see Table 1). Although the goal of the experiment was to obtain five test runs (480 responses) per syllable set on each of two days, four subjects were either unable or unwilling to complete the entire protocol, and thus only partial data are available for them.

Consonant confusion matrices were constructed for each listener on each syllable set. Subsequent analyses were based on these individual confusion matrices. Copies of these matrices can be obtained from the authors.

RESULTS

Mean scores (percent correct) are presented in Table 3 as a function of syllable set and test run for the 19 subjects who completed both full days of

TABLE 3. Mean scores (% correct) for test runs and syllable sets.

Syllable Set	*Day*	*Test Run* 1	2	3	4	5	*Mean*
CV-1	1st	41.00	43.66	45.93	46.72	47.17	44.90
	2nd	44.87	46.59	46.90	47.98	48.24	46.92
VC-1	1st	50.66	54.95	55.84	54.88	56.33	54.53
	2nd	57.82	60.05	62.74	61.02	62.03	60.73
CV-2	1st	52.09	56.29	58.74	60.02	60.90	57.61
	2nd	61.08	62.31	63.14	62.16	63.81	62.50
VC-2	1st	48.56	50.76	53.13	52.40	54.63	51.90
	2nd	55.01	59.01	59.06	58.00	61.11	58.44

testing. Several aspects of Table 3 require comment. First, a comparison of these consonant-identification scores with the W-22 scores in Table 1 replicates the well-known result that nonsense syllables are more difficult to identify than monosyllabic words (Hirsh, Reynolds, and Joseph, 1954). Second, the comparison of means for syllable sets CV-1 and VC-1, which contain the same 16 consonants, indicates that those consonants were more identifiable in the VC than in the CV context. This result, which replicates our earlier finding for normal-hearing subjects (Wang and Bilger, 1973), appears to be in conflict with the generally accepted finding that initial consonants are more identifiable than final consonants (Owens et al., 1972). While the present result may be specific to the talker used here, we would add that the studies that find initial consonants more identifiable than final consonants have consistently used monosyllabic words, primarily CVC in form, as stimuli. We would suggest that this apparent disparity is not an inconsistency but evidence of the difficulty of generalizing from CVC monosyllabic words to VC nonsense syllables. For example, the final consonant in CVC words is often poorly articulated (Silverman and Hirsh, 1956) and its identification can be based on linguistic structure rather than acoustic energy. Also in the present context, we suspect that the vowel in VCs provided an alerting signal not present in the CV context.

Stability of Performance over Time and Syllable Sets

Although we were not primarily interested in each subject's overall level of performance on the four syllable sets, since feedback was provided throughout the experiment, we wished to determine whether there were any learning or practice effects. Significant effects were obtained for syllable sets ($F = 14.39$, $df = 3{,}54$), days ($F = 40.52$, $df = 1{,}18$), and test runs ($F = 21.51$, $df = 8{,}144$), $p < 0.01$ in all cases.

Reliability of performance across syllable sets and time was also assessed. From the analysis of variance an estimate of reliability across syllable sets was obtained, $\alpha = 0.890$ (Cronbach, 1971). Test-retest correlations for the four syllable sets were: CV-1, 0.979; VC-1, 0.960; CV-2, 0.985; and VC-2, 0.960. Clearly, from a psychometric standpoint, individual differences between subjects on a consonant discrimination task are highly reliable.

The stability of the consonant confusion matrices for individual subjects, however, is of most direct concern here. To evaluate the stability of individual subjects' confusion matrices, two correlations were calculated. In the first correlation, the entries on the main diagonal of the confusion matrix for one day were correlated with those for the second day to determine the extent to which the relative difficulty of consonant phonemes is predictable from one occasion to the next. The second correlation utilized all of the cells of the confusion matrix to determine the extent to which the frequencies of both correct responses and specific confusions are predictable from one occasion to the next. (These correlations are analogous to test-retest reliability coefficients, but they cannot be treated as reliabilities, because they were computed within a single subject and not across a group of subjects.) The results of these correlational analyses are summarized in Table 4. Both correlational measures

TABLE 4. Stability coefficients for consonant confusion matrices.

Type of Coefficient	*Syllable Set*			
	CV-1	*VC-1*	*CV-2*	*VC-2*
Diagonal				
Median	0.910	0.832	0.842	0.850
Range	0.425-0.970	0.577-0.970	0.585-0.955	0.587-0.959
Matrix				
Median	0.937	0.928	0.925	0.942
Range	0.609-0.979	0.594-0.982	0.622-0.992	0.619-0.979

suggest that the present data were highly stable, although the matrix correlations are inflated to some extent by the large number of cells with zero frequencies. (Because the distribution of frequencies over all 256 cells of the confusion matrix was clearly not normal, the correlations were not tested for significance. They are reported only as descriptive statistics.)

From the foregoing analyses it is clear that the consonant-identification task produces highly stable patterns of performance in individual subjects. Except for purposes of auditory rehabilitation, however, prediction of specific consonant confusions is of little interest. Rather, it would seem desirable to employ a data reduction technique which permits both a qualitative and a quantitative comparison of confusion matrices for individual subjects.

To accomplish this we have analyzed the data in terms of transmitted information, using a sequential information analysis (SINFA) to describe patterns of phonemic confusions in terms of phonological and articulatory features that we have described previously (Wang and Bilger, 1973). There we pointed out that the use of phonological and articulatory features to describe phonemic confusions does not imply that such features can be considered to be perceptual constructs. This is because such features do not show sufficient congruence across listening conditions and syllable sets to support their status as perceptual constructs. However, they can provide a ready basis for summarization of confusions and it is in this sense that they are employed here.

Feature Analysis of Confusions

The features used to analyze the present data are identified with respect to the consonants used in Table 5. Essentially these features are those described

TABLE 5. Features specified for the sequential information analysis.

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
h^w	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

by Miller and Nicely (1955) and Chomsky and Halle (1968). The place feature we used, however, was Wickelgren's (1966) five-valued extension of the Miller and Nicely place feature. In addition to their features, the feature of sibilance is also included, because previous perceptual studies with normal-hearing subjects have indicated that it is particularly well perceived (Singh, Woods, and Tishman, 1972; Singh, Woods, and Becker, 1973; Singh and Singh, 1972; Weiner and Singh, 1974; Wang and Bilger, 1973).

It is well known that the stimulus features themselves (see Table 5) are not independent of one another. Redundancy in the stimulus features complicates the interpretation of performance on specific features, because it is not known whether performance on a particular feature may be attributed to performance on another closely related feature. To circumvent this problem we have analyzed the confusion matrices using sequential information analyses (SINFA) that we have described elsewhere (Wang and Bilger, 1973).

Input to the analysis consists of the stimulus-response confusion matrix and a set of stimulus features. In the first iteration, the feature with the highest percentage of information transmitted is identified. In the second iteration the effects of this feature are partialed out and the remaining features compared. The feature with the highest percentage of conditional information transmitted is then identified. In the third iteration, the effects of both features previously identified are partialed out and the remaining features again compared. Iterations continue in this manner until one of three termination criteria is met. Output from the analysis consists of an ordered set of stimulus features and the (conditional) transmitted information associated with each. Since the effects of feature redundancy are partialed out in each iteration, the sum of the conditional transmitted information values represents the amount of transmitted information accounted for by the stimulus features (Wang and Bilger, 1973).

An illustration of SINFA for Subject 1, CV-1 syllable set, is given in Table 6. Entries for the first iteration show that the amount of information available for transmission varies with the feature. The three-category place-of-articulation feature presents 1.561 bits of information, whereas the feature back presents only 0.544 bits. The amount of feature information transmitted is shown in the second column, and is expressed as a proportion of the available information in the third column. Relative to the other features, sibilance is very well perceived, 71.8% of the available information having been transmitted. Accordingly, sibilance is identified as the most important feature in the first iteration.

Entries for the second iteration indicate that the effect of holding sibilance constant is to reduce the available information for the other features. Stridence, frication, and duration are especially affected. The reduction in feature information is equal to the redundancy of a given feature and the feature of sibilance. The second column gives the amount of conditional information transmitted, and the third column gives the proportion of conditional information transmitted. It may be noted that the effect of holding sibilance constant is to reduce the proportion of information transmitted for some features and to increase it for others. This suggests that feature redundancies can either amplify or attenuate the apparent performance level associated with other features. Since entries in the third column indicate that voicing has the highest proportion of conditional information transmitted, it is the feature identified in the second iteration.

Entries for the remaining iterations are analogous to those for the second. In the third iteration, the highest performance level is associated with the feature duration, and in the fifth, with the feature high/anterior. In the fourth iteration, the two features continuant and frication are indistinguishable, and the highest performance level is associated with them. Similarly, in the sixth

TABLE 6. Sequential information analysis of consonant confusions for CV-1 syllable set, Subject 1.

	Feature	Feat. Inf.	Trans. Inf.	Prop. Trans. Inf.
Iteration 1	Hi./An.	0.954	0.462	0.484
	Back	0.544	0.050	0.092
	Coron.	0.955	0.200	0.209
	Voice	1.000	0.612	0.612
	Cont.	1.000	0.368	0.368
	Strid.	1.000	0.296	0.296
	Fric.	0.955	0.332	0.348
	Dur.	0.812	0.516	0.636
	Place	1.561	0.568	0.364
	Sibil.	0.954	0.685	0.718
	Feature	*Cond. Feat. Inf.*	*Cond. Trans. Inf.*	*Prop. Cond. Trans. Inf.*
Iteration 2; Constant: Sibil.	Hi./An.	0.740	0.318	0.429
	Back	0.409	0.051	0.124
	Coron.	0.597	0.055	0.092
	Voice	0.979	0.637	0.651
	Cont.	0.940	0.395	0.420
	Strid.	0.447	0.081	0.180
	Fric.	0.597	0.212	0.356
	Dur.	0.343	0.174	0.507
	Place	1.240	0.370	0.298
Iteration 3; Constant: Sibil., Voice	Hi./An.	0.703	0.325	0.462
	Back	0.381	0.054	0.140
	Coron.	0.582	0.061	0.105
	Cont.	0.906	0.459	0.506
	Strid.	0.443	0.111	0.252
	Fric.	0.574	0.277	0.483
	Dur.	0.333	0.182	0.546
	Place	1.197	0.405	0.338
Iteration 4; Constant: Sibil., Voice, Dur.	Hi./An.	0.576	0.209	0.363
	Back	0.381	0.054	0.140
	Coron.	0.582	0.061	0.105
	Cont.	0.574	0.277	0.483
	Strid.	0.443	0.111	0.252
	Fric.	0.574	0.277	0.483
	Place	1.070	0.290	0.270
Iteration 5; Constant: Sibil., Voice, Dur., Cont.	Hi./An.	0.487	0.180	0.370
	Back	0.292	0.027	0.093
	Coron.	0.544	0.091	0.167
	Strid.	0.238	0.005	0.019
	Place	0.924	0.284	0.307
Iteration 6; Constant: Sibil., Voice, Dur., Cont., Hi./An.	Coron.	0.431	0.085	0.198
	Strid.	0.234	0.005	0.020
	Place	0.431	0.085	0.198

iteration, place and coronal are indistinguishable and have the highest performance level. In both iterations the pair of features is perfectly redundant with respect to one another and the choice of a single feature name to represent the result of that iteration is arbitrary.

SINFA was performed on each confusion matrix for each subject. Since the matrices for the 19 subjects who completed the entire experiment were shown to be highly stable, matrices for the remaining four subjects, on whom only partial data were available, were analyzed also. SINFA summaries for all subjects on each syllable set are given in the Appendix.

The SINFA summaries presented in the Appendix support two observations about patients' consonant confusions. First, subjects with approximately the same overall level of performance may achieve that level by different means. Second, subjects with different levels of performance may show similar feature profiles.

Consider Subjects 1 and 10. Although the percentage of information transmitted is about the same for both subjects, the stimulus features which are relatively well perceived are different. Subject 1 identifies sibilance, duration, nasality, and voicing very well, whereas Subject 10 shows more variability across syllable sets and does well on features such as anterior, back, and frication, in addition to voicing and nasality. Subject 11 resembles Subject 1 in her relatively good perception of sibilance, duration, nasality, and voicing, although the percentage of information transmitted is lower for Subject 11 than for Subject 1.

Scaling of Intersubject Similarity

Since a major goal of this study was to determine whether subjects with sensorineural hearing loss form homogeneous subgroups on the basis of their consonant confusion patterns, we quantified the degree of similarity between individual subjects and subjected the similarities to a pairwise multidimensional scaling analysis (Johnson, 1973).

The similarity metric was based, therefore, on the results of the sequential information analysis. The results of a single analysis were coded as a vector of weights for each of the stimulus features. The feature identified in the first iteration received the highest weight; the feature identified in the last iteration received the lowest weight; and the features not identified in the analysis received zero weight. Since the average number of features identified varied somewhat with syllable sets (CV-1, 5.17; VC-1, 5.96; CV-2, 7.70; and VC-2, 6.65), the maximum weight assigned varied from five to eight across the four sets. Whenever the number of features identified exceeded the maximum weight, the lowest ranking features were all assigned weights of one.

The similarity between any two subjects was defined as the sum of the products of corresponding feature weights. This derived proximity measure (Shepard, 1972, p. 24) is closely related to the correlation between two sets of weights. For example, for Subject 1, CV-1 set, sibilance, voicing, duration, continuance, high/anterior, and coronal received weights from five to one respectively; for Subject 10, high/anterior, frication, voicing, and place received weights from five to two; and for Subject 11, voicing, sibilance, continuance, high/anterior, and place received weights from five to one. The resulting similarity measures for these three subjects were as follows: 1 and 10, 17; 1 and 11, 48; and 10 and 11, 27. The effect of calculating the similarities in this way was to give greatest weight to the results of the earlier iterations and to allow all features which two subjects had in common to contribute to the measure. It is especially desirable to give low weight to features identified in the later iterations, since, as noted earlier, the feature names utilized in these instances are often arbitrary.

Intersubject similarity was calculated for each of the 253 possible comparisons in each syllable set. In order to assess the statistical significance of the similarity score for a pair of subjects, however, it was necessary to generate the distribution of all possible scores resulting from a random sampling and ordering of features. For example, since the mean number of features identified for the CV-1 set was 5.17, we generated the distribution of all possible scores resulting from sampling five from a set of 10 features and assigning a maximum weight of five to the first feature sampled. As can be seen in Table 7, this distribution has a mean of 26.4, a standard deviation of

TABLE 7. Parameters of the theoretical sampling distribution of the similarity metric for each syllable set. The critical similarity is defined as the similarity at the 95th percentile of the sampling distribution.

Syllable Set	*Number of Features*	*Features Sampled*	*Maximum Weight*	*Mean*	*SD*	*Critical Similarity*
CV-1	10	5	5	26.4	9.8	43
VC-1	10	6	6	47.4	14.4	72
CV-2	16	8	8	99.3	28.8	148
VC-2	11	7	7	75.1	20.3	109

9.8, and scores greater than 43 occur less than 5% of the time. The parameters of the sampling distributions for the remaining syllable sets are also shown in Table 7.

Since a similarity score greater than the critical value shown in Table 7 occurs with $p < 0.05$ on the basis of random sampling, scores greater than the critical value were considered indicative of greater than chance similarity between subjects. Out of 253 possible comparisons for each syllable set, the percentage of scores exceeding the critical value was: CV-1, 39.5%; VC-1, 14.6%; CV-2, 26.5%; and VC-2, 21.3%. Since these figures are well above the 5% significant comparisons which would be expected on the basis of chance, it was concluded that there were reliable similarities in consonant confusion patterns for some pairs of subjects.

A pairwise nonmetric multidimensional scaling procedure (Johnson, 1973) was used to extract the patterns of intersubject similarity. We were not interested in using scaling to redefine the perceptual dimensions of the consonants. The features used in the information analyses are quite compatible with the stimulus dimensions other investigators have found using multidimensional scaling methods. Given that we were primarily interested in the similarities between the subjects, Johnson's (1973) procedure was chosen as an efficient method of deriving a subject space. The analysis was carried out separately for each syllable set, and in addition, an analysis was performed on similarities pooled across the four sets. The pooled similarity matrix was obtained by converting the similarity measures for each set to standard scores and summing across sets for each pair of subjects.

The lack-of-fit measure for this analysis, Θ, reflects the extent to which the rank order of original similarities matches inversely the rank order of interpoint distances in the scaling solution. The measure is similar to Kruskal's (1964) $stress_2$ in that it varies from 0 to 1.0. Johnson (1973) conducted a small simulation study to determine the expected value of Θ with random input. He found that Θ varies with the number of subjects and the number of dimensions in the solution. We therefore conducted a simulation study of our own to determine the expected value of Θ with randomly generated similarities for 23 subjects and a two-dimensional solution. For 10 simulations with random input, Θ varied from 0.387 to 0.495, with a median value of 0.461. Since the values of Θ obtained from our data were all considerably smaller than this, we accepted solutions in two dimensions as adequate. Although Θ appeared to decrease significantly with a three-dimensional solution for the two VC syllable sets, the two-dimensional solution appeared optimal for the CV syllable sets. Since our aim was to cluster the subjects rather than to identify or interpret the dimensions arising from the scaling analysis, only the two-dimensional configurations are reported here.

The results of the scaling analysis for the pooled data from all syllable sets are presented in Figure 1. Since this configuration was based on all data from the experiment, we used it as a point of departure for determining whether subjects tend to form subgroups on the basis of their consonant confusions. Inspection of the configuration suggested that three clusters of points could be identified. One cluster consisted of six points with relatively low weights on both dimensions. The remaining two clusters were differentiated by their weights on Dimension 2.

Since the scaling solution is unaffected by rotation or translation of the coordinate axes, the four configurations initially obtained for the individual syllable sets were modified, if necessary, to increase their similarity to Figure 1. The resulting configurations are shown in Figures 2-5. In each of the five figures, points which tend to form clusters have been enclosed, and identification of clusters for Figures 2-5 was guided by those identified in Figure 1.

Inspection of the figures reveals that 13 subjects can be grouped consistently for all five analyses; seven subjects can be grouped consistently for four out of five analyses; and three subjects can be grouped consistently for three out of five analyses. Therefore, out of 115 classifications, only 14 could be considered misclassifications. These misclassifications are indicated in the figures by underlining. In no case was a single subject classified in more than two groups. Nearly all misclassifications (12 out of 14) occurred for CV syllable sets and all but one involved misclassification into or out of Group A. Before

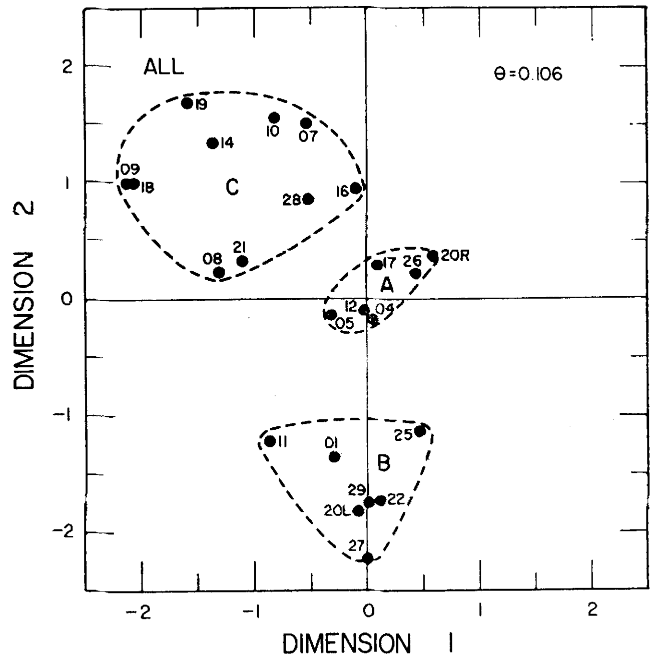

FIGURE 1. Scaling solution of consonant confusions based on intersubject similarities pooled across all four syllable sets. The numbers shown here represent individual subjects and are used consistently in this and subsequent figures. For Subject 20, data for both the normal (20R) and the impaired ear (20L) are included. The lack-of-fit measure, θ, appears in the upper-right-hand corner of this and subsequent figures.

any attempt is made to determine if there are audiometric correlates of membership in these three groups, let us review the constituency of each group and summarize their performance in terms of the feature analysis we conducted.

Group A consists of Subjects 4, 5, 12, 17, 20R, and 26. Sibilance is a well-perceived feature for these subjects. Nasality and high/anterior also are well perceived, although for two subjects, back rather than high/anterior is identified in the VC-2 syllable set. For the CV-1 and VC-1 sets, frication tends to be identified by the third or fourth iteration, as does voicing. Voicing is also identified in the VC-2 set, but typically in the very latest iterations.

Group B consists of Subjects 1, 11, 20L, 22, 25, 27, and 29. In terms of their performance in the consonant-identification task these subjects are characterized by consistently good identification of the features sibilance, duration, and voicing, generally in that order. The feature high/anterior, or a closely related place-of-articulation feature, is consistently identified in later iterations for all syllable sets. The feature continuance is identified in later iterations for the CV-1 and VC-1 sets only. Finally, it is noteworthy that the feature nasality is not well identified and is identified late in the analyses if at all.

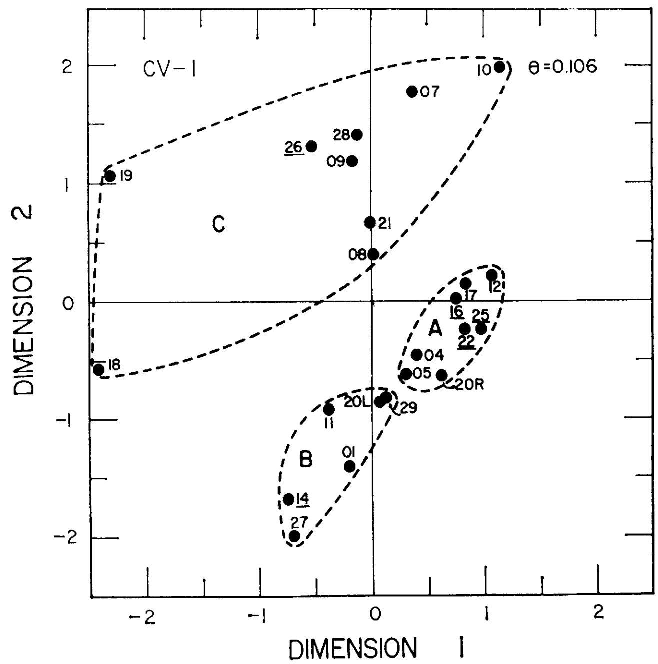

FIGURE 2. Scaling solution of consonant confusions based on intersubject similarities for the CV-1 syllable set. Underlined subject numbers indicate that a subject was misclassified for this syllable set.

Group C consists of Subjects 7, 8, 9, 10, 14, 16, 18, 19, 21, and 28. Since this is the largest group, and since it contains the two subjects (16 and 28) who were classified least consistently, it is not surprising that there is somewhat more heterogeneity in this group. Moreover, the differences between individual subjects are most apparent in the CV-2 set which contains the largest number of features and the highest degree of feature redundancy. Perhaps the most notable characteristic of these subjects is their inability to identify sibilance. It is this which sets them apart from the other groups most clearly. Not only is sibilance not identified early in the analysis, it is generally not identified at all. Interestingly, four out of the five misclassifications of subjects in this group involve the presence of sibilance. Relative to the other groups, Group C perceives nasality very well. Voicing is prominent in those sets where nasality is not distinctive and is less well perceived than nasality when the latter is distinctive. With these exceptions, the performance of Group C is otherwise quite similar to that of Group A.

Although the three groups of subjects described above do not differ radically, there are sufficiently consistent differences in their performance on the consonant identification task to permit successful classification into subgroups. Since this is the case, we wished to determine whether membership in the various groups might be related to other characteristics of the patients. In

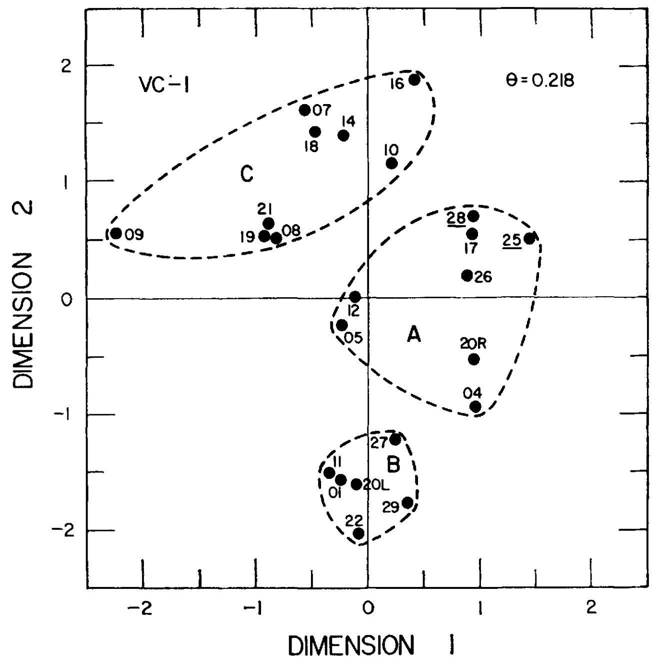

FIGURE 3. Scaling solution of consonant confusions based on intersubject similarities for the VC-1 syllable set. Underlined subject numbers indicate that a subject was misclassified for this syllable set.

particular we were interested in using audiometric test data to predict group membership.

A summary of the audiometric data for the three groups of subjects is presented in Table 8. Group A contains the three subjects designated as normal controls, 4, 5, and 20R. The remaining four subjects have mild pure-tone hearing losses ($\leq$35 dB HL) up to 1000 Hz. For frequencies above 1000 Hz, there is more heterogeneity among the subjects (Table 1). As a group, these subjects have normal or very slightly elevated speech reception thresholds and normal or slightly depressed speech discrimination scores. Subjects in Group B have moderate to severe hearing losses for pure tones and the audiogram is generally flat or rising. The speech reception threshold is elevated and the word-recognition score, W-22, is the lowest of the three groups. The subjects in Group C may be characterized as having high-frequency hearing losses. For some subjects there is no loss at the lower frequencies; for others there is a substantial loss at the lower frequencies and an even greater loss at the high frequencies. With respect to speech reception thresholds and word-recognition scores this group falls between Groups A and B, although there is considerable variability within the group.

Although the comparisons above suggest that patterns of consonant confusion can be reliably related to characteristics of a patient's audiogram, it is

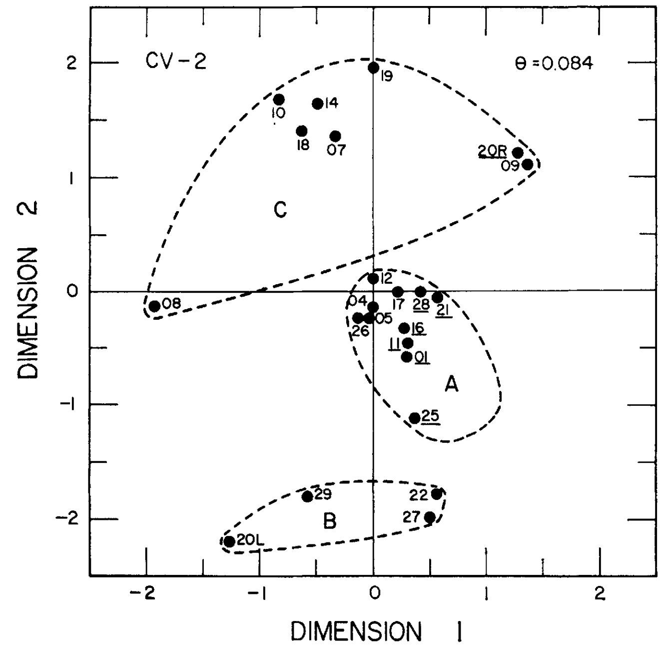

FIGURE 4. Scaling solution of consonant confusions based on intersubject similarities for the CV-2 syllable set. Underlined subject numbers indicate that a subject was misclassified for this syllable set.

TABLE 8. Averaged audiometric characteristics of three subgroups of subjects.

	Pure-Tone Thresholds (ANSI, 1969)						*Speech*	
Group	*250 Hz*	*500 Hz*	*1 kHz*	*2 kHz*	*4 kHz*	*8 kHz*	*SRT*	*W-22*
A	16.7 dB	15.8 dB	14.2 dB	17.5 dB	25.8 dB	25.0 dB	10.3 dB	94.3 dB
B	48.6	50.0	51.4	56.4	51.4	52.9	42.7	61.4
C	20.0	23.0	32.5	52.5	78.0	69.0	22.2	82.2

important to bear in mind that the grouping of subjects constitutes a post hoc analysis of the data and that the averaging of audiometric profiles tends to obscure individual differences (and inconsistencies) among subjects. The conclusion would be strengthened if it could be shown that the group membership of an independent sample of subjects could then be predicted from audiometric data alone. Fortunately, relevant data are available. As part of another study, Reed (1975) gathered consonant confusion data, for the VC-1 syllable set, on a sample of 12 subjects using the same equipment and procedures as in the present study. Descriptive and audiometric data for these subjects are presented in Table 9, together with the predicted group membership of each subject.

A sequential information analysis was performed on the confusion matrix

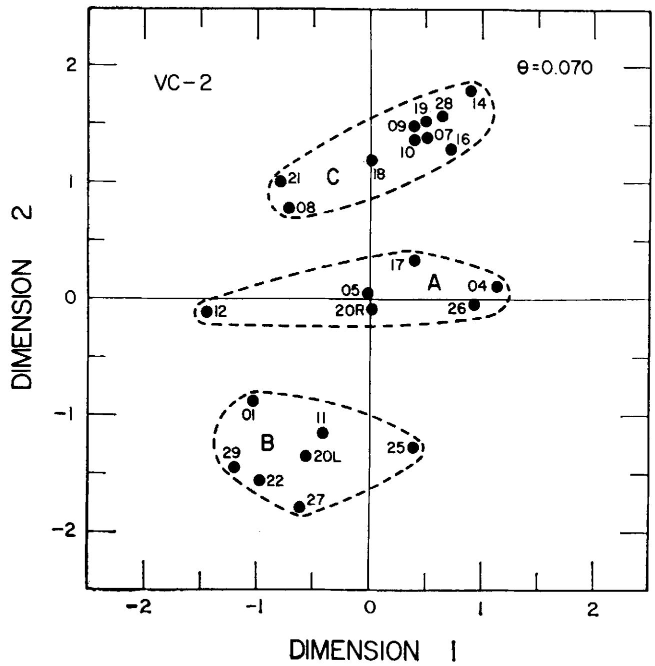

FIGURE 5. Scaling solution of consonant confusions based on intersubject similarities for the VC-2 syllable set. Underlined subject numbers indicate that a subject was misclassified for this syllable set.

TABLE 9. Descriptive and audiometric data for additional subjects and prediction of group membership.

			Pure-Tone Thresholds (ANSI, 1969)						*Speech*	*Pred.*	
Subject	*Age*	*Sex*	*250 Hz*	*500 Hz*	*1 kHz*	*2 kHz*	*4 kHz*	*8 kHz*	*SRT*	*W-22*	*Group*
AL	46	M	60 dB	65 dB	60 dB	60 dB	70 dB	65 dB	60 dB	54%	B
B	42	M	60	50	60	50	60	30	42	84	B
C	41	F	45	45	45	45	55	35	38	72	B
D	45	M	60	55	55	45	50	45	50	64	B
E	30	F	65	65	55	50	20	10	38	52	B
F	33	M	55	65	65	60	80	75	66	48	B
Mean	39.5		57.5	57.5	56.7	51.7	55.8	43.3	49.0	62.3	
AR	46	M	10	5	10	10	10	0	8	96	A
G	50	F	15	10	5	5	5	15	0	100	A
H	24	F	10	5	10	10	5	0	10	100	A
Mean	40.0		11.7	6.7	8.3	8.3	6.7	5.0	6.0	98.7	
I	51	M	5	15	70	110	110	90	40	20	C
J	49	F	15	25	75	75	80	70	48	48	C
K	49	M	0	20	95	110	110	90	24	20	C
Mean	49.7		6.7	20.0	80.0	98.3	100.0	83.3	37.3	29.3	

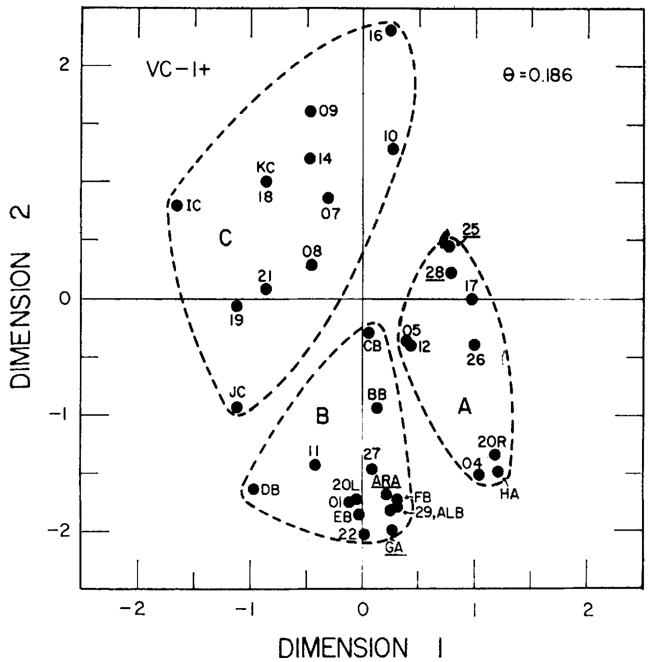

FIGURE 6. Rescaling of VC-1 consonant set based on our data plus the subjects of Reed (1975). Reed's subjects are identified by two upper-case letters, the first identifying the subject (A-K), and the second identifying predicted group membership (A, B, or C). For Subject A, data for both the normal (ARA) and the impaired ear (ALB) are included.

for each subject; the results were coded as a vector of feature weights; and the data pooled with those of the original 23 subjects. The resulting 595-cell similarity matrix was analyzed by the Johnson procedure. The results are shown in Figure 6.

Although the new configuration for the VC-1 set contains 35 points rather than 23, Θ is only 0.186. Because Θ would be expected to rise as points are added to the configuration (Johnson, 1973), we accepted the solution in two dimensions as quite adequate.

Generally, the predicted groupings were obtained, although the results are much clearer for Group C than for the other groups. Subject KC perfectly duplicated his previous performance and was again classified in Group C. Subjects IC and JC, who also showed severe high-frequency loss and a steep audiogram, were easily included in Group C.

Subject HA, the only new subject with bilaterally normal hearing, showed performance highly similar to that of two other normal controls, Subjects 4 and 5. Two subjects who were predicted to fall in Group A, however, clearly belonged in Group B. It is interesting that both of these subjects, GA and ARA, had a flat hearing loss in the opposite ear. Although the test ear was normal, audiometric data for the opposite ear would have led to a prediction of membership in Group B. It may be seen that when subject ALB was tested in the poorer ear, his performance did place him in Group B as expected.

Of the remaining subjects, all were predicted to fall in Group B because of flat (or upward sloping) audiograms. Although Subjects BB and CB could be included in Group A, it was also possible to place them in Group B without unduly misrepresenting the clusters. It is clear, however, that the results for these two subjects should be considered borderline. Subjects DB, EB, and FB fell in Group B as predicted.

DISCUSSION AND CONCLUSIONS

If the consonant-identification task is considered a test differentiating listeners in terms of performance level, it is clear that individual differences between subjects are highly reliable over time and also over different sets of test materials. Although the level of performance may not be well predicted from audiometric data, it is clearly predictable from a limited sample of discrimination responses.

Although reliability coefficients emphasize the stability of individual differences between listeners, we also found that subjects improve consistently with practice on the recognition task. This was true even when the test-retest interval was greater than one year. The implications of this finding for auditory rehabilitation need to be explored more fully. Specifically, it is necessary to determine the extent to which this improvement reflects increased familiarity with an artificial laboratory task, and the extent to which it reflects improvement in speech-recognition ability which will generalize to performance outside the laboratory.

When the consonant confusions of individual subjects were examined for consistency over time, it was found that they were highly stable. Both the relative difficulty of individual consonants and the relative frequency of specific consonant confusions appear to be highly reliable. Similar findings by Lawrence and Byers (1969), based on four voiceless fricatives, may thus be safely generalized to all English consonants. This suggests that it should be possible to describe the nature of a single patient's speech-recognition problem in some detail, and, possibly, to prescribe individualized auditory rehabilitation training.

When consonant confusion patterns are described in terms of feature identification, two general findings emerge. First, similar patterns of feature perception may be observed in patients with very different levels of performance. The comparison of Subjects 1 and 11 discussed above is an example of this. Even more dramatic examples are provided by Subjects 9 and 26 for the CV-1 set; Subjects 27 and 29 for the VC-1 set; Subjects 7 and 18 for the CV-2 set; and Subjects 7 and 9 for the VC-2 sets (see Appendix). Performance levels, defined in terms of the percentage of information transmitted, vary by as much as 56.4% between members of these pairs, although relative importance of the various features is highly similar. Second, subjects performing at roughly the same level do not necessarily show similar patterns of confusion. Subjects 9 and 27, for example, perform very poorly on all four syllable sets, yet they show reliably different patterns of perceptual confusion; Subjects 7 and 20R perform very well on all four sets and also show different patterns of confusion. We can can thus conclude that the number of errors a patient makes, and the types of errors made, reflects relatively independent aspects of auditory functioning.

Of greater interest, however, is the finding that intersubject similarities in patterns of perceptual confusion are systematically related to the subjects' audiometric configurations. Owens et al. (1972) also demonstrated that groups of patients with different audiometric configurations experience different degrees of difficulty with certain consonants. Rather than group the patients on the basis of their audiograms at the outset, however, we chose the alternative approach of describing patterns of perceptual confusions for individuals, and determining, by means of multidimensional scaling, whether they tend to form homogeneous groups. The results clearly suggest that such groups do exist, and that group membership, with very few exceptions, is independent of the specific stimulus set used to test recognition. Whether group members have certain audiometric characteristics in common is an independent question. Inspection of the audiograms of the three groups suggested that they could be described as normal listeners or listeners with mild, flat losses (Group A), listeners with moderate to severe flat losses (Group B), and listeners with high-frequency losses (Group C). This conclusion was strengthened by a cross-validation in which the group membership of 10 out of 12 new patients was predicted on the basis of audiometric data alone.

Although we have emphasized the audiometric similarity of subjects within Groups A, B, and C, these groups are not perfectly homogeneous, and the performance of a few subjects would not have been well predicted on the basis of their audiometric configurations. Although this does not alter our general conclusion that audiometric configuration and consonant confusions are related, it does suggest that there are meaningful differences in consonant confusions between some subjects with similar audiograms. It is these differences, we believe, which make at least a two-dimensional scaling solution necessary. One dimension appears to be sufficient for broadly differentiating among different groups of listeners. Dimension 2 in Figures 1-5 generally places "normal" listeners in the center of the configuration between the two more severely impaired groups. A second dimension is required in order to represent the residual differences between listeners within these groups. For this reason we have not attempted to label or further interpret the dimensions obtained from the scaling analyses.

Our findings agree with those of Owens et al. (1972), in that listeners with high-frequency hearing loss had difficulty with sibilant consonants. With respect to consonant confusion, however, our results are in direct opposition to theirs. Owens et al. found that the specific errors which were most likely to occur were independent of audiometric configuration. Differences revealed by the feature analysis in the present study, on the other hand, suggest that there is a relationship between audiometric configuration and pattern of consonant confusions. The most plausible explanation of the discrepancy between the two studies lies in the nature of the recognition task. Owens et al. obtained confusions using a multiple-choice word-recognition task. In the present study the stimuli were nonsense syllables and there were 16 alternatives on each trial. Moreover, the number of trials per subject was greater in the present study. Thus, the task used here was probably more sensitive to different rates of phonemic confusion than the task used by Owens et al.

Additional evidence concerning patterns of phonemic perception has recently been reported by Walden and Montgomery (1975). Three groups of hearing-impaired listeners judged the similarity of 190 pairs of CV syllables formed by combining 20 consonants with the vowel /a/. The similarity judgments were analyzed using the individual differences scaling analysis, INDSCAL. Walden and Montgomery found that the three groups of listeners (normal hearing, high-frequency loss, and flat loss) were discernible in the three-dimensional space obtained from INDSCAL. Moreover, the stimulus dimensions found to be most important for the three groups considered individually are consistent with the features found to be most important for comparable groups in our study. The only exception to this was their finding that sibilance was a relatively important dimension for listeners with high-frequency loss. Given that the two studies were based both on a different response task and a different type of analysis, detailed comparisons are both difficult to make and to interpret. What is more important, however, is that despite the methodological differences between the studies, they provide converging evidence for the conclusion that patterns of consonant perception in patients with sensorineural hearing loss may be predicted from audiometric data.

ACKNOWLEDGMENT

This research was supported by grants from NINCDS to the University of Pittsburgh (NS 07790) and to the Eye and Ear Hospital of Pittsburgh (NS 12501) and in part by a grant from SRS (14-P-55126). In addition, much of the data analysis was also supported by faculty research funds made available to M. D. Wang through the Social Sciences Statistical 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 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; Nancy R. Stenson and Norma T. Hopkinson for assistance in selecting and scheduling subjects for the experiments; Linet Feigel and Walt Jesteadt for their assistance in programming and data analysis; and Louis Braida and Charlotte Reed for their helpful comments on an earlier version of this manuscript. Special thanks also are due to the staff of the Social Sciences Statistical Center of UMBC and to Edward Fahrmeier, who called our attention to the particular scaling procedure used here. Requests for reprints or individual matrices should be addressed to Robert C. Bilger, Bioacoustics Laboratory, Department of Otolaryngology, Eye and Ear Hospital, 230 Lothrop Street, Pittsburgh, Pennsylvania 15213; or Marilyn D. Wang, Department of Psychology, UMBC, 5401 Wilkens Avenue, Baltimore, Maryland 21228.

REFERENCES

CHOMSKY, N., and HALLE, M., *The Sound Pattern of English.* New York: Harper and Row (1968).

CRONBACH, L., *Essentials of Psychological Testing.* (3rd ed.) New York: Harper and Row (1971).

HIRSH, I. J., REYNOLDS, E. G., and JOSEPH, M., Intelligibility of different speech materials. *J. acoust. Soc. Am.*, **26**, 530-538 (1954).

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

KRUSKAL, J. B., Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika,* **29**, 1-27 (1964).

LAWRENCE, D. L., and BYERS, V. W., Identification of voiceless fricatives by high frequency hearing impaired listeners. *J. Speech Hearing Res.*, **12**, 426-434 (1969).

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., and SCHUBERT, E. D., The development of consonant items for speech discrimination testing. *J. Speech Hearing Res.*, **11**, 656-667 (1968).

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

OYER, H., and DOUDNA, M., Structural analysis of word responses made by hard of hearing subjects on a discrimination test. *Arch Otolaryng.*, **70**, 357-364 (1959).

REED, C. M., Identification and discrimination of vowel-consonant syllables in listeners with sensorineural hearing loss. *J. Speech Hearing Res.*, **18**, 773-794 (1975).

SCHULTZ, M. C., Suggested improvements in speech discrimination testing. *J. aud. Res.*, **4**, 1-14 (1964).

SHEPHARD, R. N., A taxonomy of some principle types of data and of multidimensional methods for their analysis. In R. N. Shephard, A. K. Romney, and S. B. Nerlove (Eds.), *Multidimensional Scaling Theory and Applications in the Behavioral Sciences.* Vol. 1. New York: Seminar Press (1972).

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

SILVERMAN, S. R., and HIRSH, I. J., Problems related to the use of speech in clinical audiometry. *Ann. Otol. Rhinol. Laryng.*, **64**, 1234-1244 (1955).

SINGH, S., and SINGH, K., Search for the perceptual features of the 29 prevocalic Hindi consonants. (abstract) *J. acoust. Soc. Am.*, **52**, 112 (1972).

SINGH, S., WOODS, D. R., and BECKER, G. M., Perceptual structure of 22 prevocalic English consonants. *J. acoust. Soc. Am.*, **52**, 1698-1713 (1973).

SINGH, S., WOODS, D. R., and TISHMAN, A., An alternative MD-SCAL analysis of the Graham and House data. *J. acoust. Soc. Am.*, **51**, 666-668 (1972).

WALDEN, B. E., and MONTGOMERY, A. A., Dimensions of consonant perception in normal and hearing-impaired listeners. *J. Speech Hearing Res.*, **18**, 444-455 (1975).

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

WEINER, F. F., and SINGH, S., Multidimensional analysis of choice reaction time judgment on pairs of English fricatives. *J. exp. Psych.*, **102**, 615-620 (1974).

WICKELGREN, W. A., Distinctive features and errors in short-term memory for English consonants. *J. acoust. Soc. Am.*, **39**, 388-398 (1966).

Received September 2, 1975.
Accepted May 16, 1976.

APPENDIX

Syllable Set

	CV-1			VC-1			CV-2			VC-2		
Subject	*Iteration*	*Feature*	*Cond. Tr. Inf.*	*Iteration*	*Feature*	*Cond. Tr. Inf.*	*Iteration*	*Feature*	*Cond. Tr. Inf.*	*Iteration*	*Feature*	*Cond. Tr. Inf.*
01	1	Sibil.	0.685	1	Dur.	0.723	1	Sibil.	0.780	1	Dur.	0.631
	2	Voice	0.637	2	Sibil.	0.359	2	Dur.	0.235	2	Nasal	0.446
	3	Dur.	0.182	3	Voice	0.674	3	Nasal	0.429	3	Sibil.	0.281
	4	Cont.	0.277	4	Cont.	0.355	4	Voice	0.601	4	Voice	0.506
	5	Hi./An.	0.180	5	Hi./An.	0.298	5	Place	0.559	5	Cont.	0.296
	6	Coron.	0.085	6	Coron.	0.125	6	Strid.	0.161	6	Hi./An.	0.235
							7	Round	0.103			
Total			2.046			2.533			2.867			2.395
04	1	Sibil.	0.755	1	Sibil.	0.784	1	Dur.	0.512	1	Back	0.470
	2	Voice	0.677	2	Back	0.364	2	Nasal	0.469	2	Sibil.	0.704
	3	Hi./An.	0.557	3	Dur.	0.245	3	Sibil.	0.423	3	Nasal	0.397
	4	Fric.	0.322	4	Cont.	0.314	4	Anter.	0.522	4	Fric.	0.254
	5	Cont.	0.127	5	Coron.	0.345	5	Cons.	0.177	5	Cont.	0.211
	6	Coron.	0.195	6	Voice	0.494	6	Coron.	0.265	6	Coron.	0.184
				7	Hi./An.	0.118	7	Round	0.161	7	Voice	0.282
							8	Voice	0.496	8	Place	0.086
							9	Cont.	0.141			
Total			2.632			2.665			3.175			2.586
05	1	Sibil.	0.744	1	Hi./An.	0.719	1	Sibil.	0.728	1	Nasal	0.582
	2	Voice	0.731	2	Sibil.	0.570	2	Dur.	0.240	2	Hi./An.	0.704
	3	Fric.	0.438	3	Voice	0.611	3	Nasal	0.463	3	Sibil.	0.361
	4	Hi./An.	0.440	4	Fric.	0.322	4	Cons.	0.534	4	Fric.	0.236
	5	Coron.	0.210	5	Coron.	0.287	5	Anter.	0.191	5	Voice	0.317
	6	Cont.	0.107	6	Cont.	0.063	6	Round	0.167	6	Coron.	0.147
							7	Voice	0.557	7	Cont.	0.058
							8	Strid.	0.238			
							9	Place	0.133			
Total			2.670			2.572			3.251			2.404

APPENDIX (cont.).

	CV-1			VC-1			CV-2			VC-2		
Subject	Iteration	Feature	Cond. Tr. Inf.	Iteration	Feature	Cond. Tr. Inf.	Iteration	Feature	Cond. Tr. Inf.	Iteration	Feature	Cond. Tr. Inf.
07	1	Hi./An.	0.840	1	Fric.	0.870	1	Nasal	0.523	1	Back	0.500
	2	Back	0.290	2	Hi./An.	0.849	2	Anter.	0.795	2	Nasal	0.584
	3	Fric.	0.489	3	Coron.	0.470	3	Cons.	0.384	3	Fric.	0.428
	4	Voice	0.789	4	Voice	0.735	4	Voc.	0.410	4	Hi./An.	0.540
	5	Coron.	0.294	5	Cont.	0.159	5	Round	0.195	5	Voice	0.511
	6	Cont.	0.126	6	Strid.	0.103	6	Voice	0.551	6	Cont.	0.148
	7	Strid.	0.039				7	Dur.	0.263	7	Place	0.278
							8	Cont.	0.173	8	Strid.	0.093
							9	Coron.	0.084			
Total			2.867			3.186			3.378			3.083
08	1	Voice	0.813	1	Voice	0.773	1	Voice	0.692	1	Voice	0.764
	2	Hi./An.	0.468	2	Fric.	0.702	2	High	0.584	2	Nasal	0.389
	3	Fric.	0.398	3	Cont.	0.273	3	Back	0.124	3	Fric.	0.477
	4	Sibil.	0.101	4	Hi./An.	0.347	4	Cons.	0.294	4	Cont.	0.249
	5	Cont.	0.057	5	Dur.	0.125	5	Anter.	0.151	5	Hi./An.	0.319
	6	Coron.	0.032	6	Coron.	0.138	6	Dur.	0.236	6	Dur.	0.118
							7	Nasal	0.131	7	Coron.	0.038
							8	Strid.	0.115			
							9	Round	0.042			
Total			1.869			2.357			2.370			2.355
09	1	Voice	0.288	1	Voice	0.315	1	Nasal	0.243	1	Nasal	0.376
	2	Hi./An.	0.106	2	Cont.	0.222	2	Voc.	0.128	2	Voice	0.205
	3	Fric.	0.114	3	Place	0.209	3	Strid.	0.123	3	Fric.	0.079
	4	Dur.	0.021	4	Coron.	0.014	4	Voice	0.123	4	Back	0.028
				5	Strid.	0.014	5	Place	0.179	5	Coron.	0.026
										6	Strid.	0.016
Total			0.528			0.773			0.795			0.730

Appendix (cont.).

	CV-1			VC-1			CV-2			VC-2		
Subject	*Iteration*	*Feature*	*Cond. Tr. Inf.*	*Iteration*	*Feature*	*Cond. Tr. Inf.*	*Iteration*	*Feature*	*Cond. Tr. Inf.*	*Iteration*	*Feature*	*Cond. Tr. Inf.*
10	1	Hi./An.	0.689	1	Back	0.466	1	Anter.	0.800	1	Nasal	0.617
	2	Fric.	0.584	2	Fric.	0.544	2	Cons.	0.348	2	Fric.	0.543
	3	Voice	0.616	3	Voice	0.648	3	Strid.	0.597	3	Back	0.268
	4	Place	0.212	4	Cont.	0.274	4	Round	0.153	4	Cont.	0.272
				5	Coron.	0.390	5	Voice	0.653	5	Voice	0.439
				6	Hi./An.	0.063	6	Nasal	0.161	6	Coron.	0.240
				7	Strid.	0.028	7	Voc.	0.062	7	Hi./An.	0.086
							8	Coron.	0.167			
Total			2.102			2.413			2.940			2.473
11	1	Voice	0.476	1	Sibil.	0.663	1	Voice	0.575	1	Sibil.	0.660
	2	Sibil.	0.322	2	Voice	0.549	2	Sibil.	0.443	2	Voice	0.555
	3	Cont.	0.265	3	Dur.	0.183	3	Dur.	0.181	3	Dur.	0.205
	4	Hi./An.	0.097	4	Cont.	0.269	4	Nasal	0.177	4	Nasal	0.211
	5	Place	0.037	5	Place	0.166	5	Low	0.090	5	Cont.	0.163
							6	Voc.	0.101	6	Back	0.035
							7	Place	0.116	7	Place	0.032
							8	Strid.	0.058			
Total			1.197			1.829			1.742			1.861
12	1	Hi./An.	0.885	1	Hi./An.	0.832	1	Nasal	0.544	1	Nasal	0.620
	2	Sibil.	0.650	2	Sibil.	0.665	2	Dur.	0.518	2	Dur.	0.543
	3	Fric.	0.351	3	Voice	0.629	3	Sibil.	0.451	3	Cont.	0.311
	4	Voice	0.712	4	Fric.	0.238	4	Anter.	0.568	4	Hi./An.	0.404
	5	Coron.	0.240	5	Coron.	0.285	5	Cons.	0.225	5	Voice	0.351
	6	Cont.	0.099	6	Cont.	0.099	6	Round	0.215	6	Place	0.119
							7	Strid.	0.239			
							8	Voice	0.633			
							9	Coron.	0.150			
Total			2.937			2.747			3.543			2.348

APPENDIX (cont.).

	CV-1			VC-1			CV-2			VC-2		
Subject	*Iteration*	*Feature*	*Cond. Tr. Inf.*	*Iteration*	*Feature*	*Cond. Tr. Inf.*	*Iteration*	*Feature*	*Cond. Tr. Inf.*	*Iteration*	*Feature*	*Cond. Tr. Inf.*
14	1	Voice	0.230	1	Voice	0.458	1	Nasal	0.423	1	Fric.	0.480
	2	Sibil.	0.190	2	Fric.	0.340	2	Round	0.303	2	Nasal	0.099
	3	Cont.	0.135	3	Coron.	0.152	3	Cons.	0.269	3	Back	0.072
	4	Coron.	0.094	4	Back	0.055	4	Strid.	0.329	4	Coron.	0.056
				5	Dur.	0.049	5	Low	0.105	5	Voice	0.060
				6	Place	0.020	6	Voice	0.214	6	Strid.	0.015
							7	High	0.106	7	Place	0.036
							8	Place	0.145			
Total			0.650			1.076			1.894			0.816
16	1	Hi./An.	0.602	1	Fric.	0.459	1	Dur.	0.488	1	Fric.	0.503
	2	Sibil.	0.386	2	Back	0.213	2	Sibil.	0.455	2	Back	0.191
	3	Voice	0.363	3	Coron.	0.297	3	Nasal	0.350	3	Nasal	0.241
	4	Fric.	0.252	4	Voice	0.214	4	Coron.	0.406	4	Hi./An.	0.156
	5	Coron.	0.156	5	Place	0.085	5	Anter.	0.361	5	Voice	0.137
	6	Cont.	0.029	6	Strid.	0.046	6	Round	0.115	6	Sibil.	0.098
							7	Voice	0.263	7	Coron.	0.040
							8	Cont.	0.110			
Total			1.788			1.314			2.548			1.365
17	1	Hi./An.	0.699	1	Back	0.399	1	Nasal	0.527	1	Fric.	0.621
	2	Sibil.	0.469	2	Sibil.	0.638	2	Sibil.	0.668	2	Hi./An.	0.582
	3	Voice	0.494	3	Fric.	0.309	3	Dur.	0.273	3	Nasal	0.220
	4	Fric.	0.278	4	Hi./An.	0.200	4	Cons.	0.302	4	Sibil.	0.175
	5	Place	0.183	5	Voice	0.547	5	Coron.	0.338	5	Cont.	0.106
				6	Coron.	0.236	6	Round	0.084	6	Voice	0.243
				7	Cont.	0.092	7	Voice	0.375	7	Coron.	0.069
							8	Anter.	0.067			
							9	Cont.	0.138			
Total			2.123			2.421			2.772			2.016

Appendix (cont.).

	CV-1			VC-1			CV-2			VC-2		
Subject	Iteration	Feature	Cond. Tr. Inf.	Iteration	Feature	Cond. Tr. Inf.	Iteration	Feature	Cond. Tr. Inf.	Iteration	Feature	Cond. Tr. Inf.
18	1	Voice	0.383	1	Voice	0.599	1	Nasal	0.400	1	Nasal	0.347
	2	Cont.	0.092	2	Fric.	0.370	2	Voice	0.398	2	Voice	0.247
	3	Place	0.093	3	Cont.	0.060	3	Low	0.113	3	Fric.	0.160
	4	Strid.	0.015	4	Back	0.040	4	Anter.	0.195	4	Back	0.036
				5	Place	0.067	5	Strid.	0.137	5	Cont.	0.048
							6	Cons.	0.023	6	Place	0.030
							7	Back	0.020			
							8	Coron.	0.019			
Total			0.584			1.136			1.306			0.867
19	1	Voice	0.269	1	Voice	0.331	1	Round	0.296	1	Nasal	0.368
	2	Fric.	0.224	2	Cont.	0.314	2	Voc.	0.274	2	Fric.	0.259
	3	Place	0.224	3	Coron.	0.196	3	Nasal	0.246	3	Voice	0.199
	4	Dur.	0.025	4	Back	0.034	4	Cons.	0.191	4	Back	0.073
				5	Hi./An.	0.017	5	Low	0.064	5	Coron.	0.096
							6	Strid.	0.094	6	Strid.	0.018
							7	Anter.	0.111	7	Hi./An.	0.017
							8	Voice	0.107			
							9	Coron.	0.022			
Total			0.741			0.892			1.405			1.028
20R	1	Sibil.	0.904	1	Back	0.511	1	Coron.	0.929	1	Nasal	0.680
	2	Hi./An.	0.743	2	Dur.	0.659	2	Nasal	0.515*	2	Sibil.	0.774
	3	Voice	0.826	3	Hi./An.	0.571	3	Cont.	0.585	3	Hi./An.	0.733
	4	Dur.	0.183	4	Cont.	0.384	4	Voc.	0.165	4	Fric.	0.301
	5	Cont.	0.366	5	Voice	0.660	5	Round	0.213	5	Cont.	0.197
	6	Coron.	0.240	6	Coron.	0.322	6	Voice	0.428	6	Voice	0.591
							7	Low	0.121	7	Coron.	0.203
							8	Anter.	0.162			
Total			3.263			3.107			3.118			3.479

APPENDIX (cont.).

	CV-1			VC-1			CV-2			VC-2		
Subject	*Iteration*	*Feature*	*Cond. Tr. Inf.*	*Iteration*	*Feature*	*Cond. Tr. Inf.*	*Iteration*	*Feature*	*Cond. Tr. Inf.*	*Iteration*	*Feature*	*Cond. Tr. Inf.*
20L	1	Sibil.	0.286	1	Dur.	0.721	1	Dur.	0.287	1	Sibil.	0.837
	2	Voice	0.170	2	Sibil.	0.375	2	Sibil.	0.163	2	Dur.	0.302
	3	Hi./An.	0.133	3	Voice	0.527	3	Voice	0.167	3	Nasal	0.273
	4	Cont.	0.046	4	Hi./An.	0.352	4	High	0.071	4	Hi./An.	0.336
	5	Place	0.022	5	Cont.	0.195	5	Back	0.017	5	Cont.	0.155
				6	Coron.	0.114	6	Place	0.036	6	Voice	0.284
Total			0.657			2.285			0.741			2.186
21	1	Voice	0.528	1	Fric.	0.697	1	Nasal	0.477	1	Cont.	0.781
	2	Fric.	0.458	2	Cont.	0.344	2	Voc.	0.387	2	Fric.	0.367
	3	Hi./An.	0.541	3	Voice	0.587	3	Sibil.	0.335	3	Nasal	0.315
	4	Sibil.	0.035	4	Hi./An.	0.385	4	Dur.	0.199	4	Voice	0.421
	5	Place	0.037	5	Dur.	0.092	5	Voice	0.265	5	Hi./An.	0.339
				6	Coron.	0.125	6	Fric.	0.142	6	Dur.	0.075
							7	Anter.	0.119	7	Coron.	0.061
Total			1.599			2.230			1.926			2.358
22	1	Sibil.	0.745	1	Dur.	0.656	1	Sibil.	0.674	1	Sibil.	0.695
	2	Hi./An.	0.312	2	Sibil.	0.279	2	Dur.	0.221	2	Dur.	0.169
	3	Voice	0.224	3	Cont.	0.257	3	Voice	0.220	3	Voice	0.271
	4	Fric.	0.080	4	Hi./An.	0.236	4	Place	0.126	4	Hi./An.	0.204
	5	Place	0.042	5	Voice	0.364	5	Strid.	0.037	5	Cont.	0.146
				6	Place	0.089				6	Nasal	0.048
										7	Coron.	0.020
Total			1.404			1.881			1.278			1.552
25	1	Hi./An.	0.160	1	Sibil.	0.248	1	Sibil.	0.205	1	Sibil.	0.373
	2	Sibil.	0.097	2	Fric.	0.087	2	Dur.	0.147	2	Nasal	0.114
	3	Voice	0.095	3	Place	0.102	3	Place	0.172	3	Cont.	0.084
	4	Dur.	0.009	4	Voice	0.114	4	Voice	0.068	4	Place	0.099
							5	Round	0.026	5	Voice	0.064
							6	Nasal	0.011			
Total			0.361			0.551			0.628			0.734

APPENDIX (cont.).

	CV-1			VC-1			CV-2			VC-2		
Subject	*Iteration*	*Feature*	*Cond. Tr. Inf.*	*Iteration*	*Feature*	*Cond. Tr. Inf.*	*Iteration*	*Feature*	*Cond. Tr. Inf.*	*Iteration*	*Feature*	*Cond. Tr. Inf.*
26	1	Hi./An.	0.791	1	Sibil.	0.884	1	Dur.	0.544	1	Sibil.	0.929
	2	Dur.	0.666	2	Back	0.402	2	Sibil.	0.509	2	Back	0.415
	3	Voice	0.810	3	Hi./An.	0.297	3	Nasal	0.444	3	Nasal	0.475
	4	Cont.	0.300	4	Fric.	0.423	4	Anter.	0.501	4	Fric.	0.269
	5	Back	0.133	5	Voice	0.719	5	Cons.	0.192	5	Hi./An.	0.271
	6	Coron.	0.193	6	Cont.	0.178	6	Round	0.193	6	Cont.	0.174
				7	Coron.	0.284	7	Voice	0.626	7	Voice	0.481
							8	Strid.	0.203	8	Coron.	0.161
							9	Coron.	0.104			
Total			2.892			3.187			3.317			3.177
27	1	Sibil.	0.130	1	Sibil.	0.314	1	Sibil.	0.268	1	Sibil.	0.485
	2	Dur.	0.053	2	Dur.	0.087	2	Dur.	0.050	2	Dur.	0.091
	3	Voice	0.096	3	Voice	0.099	3	Voice	0.062	3	Voice	0.069
	4	Cont.	0.012	4	Cont.	0.043	4	Place	0.027	4	Place	0.034
	5	Place	0.033	5	Back	0.016	5	Cont.	0.015	5	Nasal	0.013
				6	Coron.	0.015						
Total			0.323			0.573			0.422			0.691
28	1	Hi./An.	0.563	1	Back	0.304	1	Nasal	0.498	1	Nasal	0.562
	2	Voice	0.473	2	Sibil.	0.408	2	Voc.	0.288	2	Fric.	0.530
	3	Fric.	0.271	3	Fric.	0.198	3	Coron.	0.431	3	Back	0.206
	4	Place	0.134	4	Voice	0.412	4	Dur.	0.195	4	Place	0.467
	5	Strid.	0.021	5	Coron.	0.189	5	Voice	0.306	5	Cont.	0.105
				6	Cont.	0.108	6	Back	0.094	6	Voice	0.257
				7	Place	0.041	7	High	0.073			
							8	Anter.	0.112			
							9	Cont.	0.044			
Total			1.461			1.659			2.043			2.127

APPENDIX (cont.).

	CV-1			VC-1			CV-2			VC-2		
Subject	*Iteration*	*Feature*	*Cond. Tr. Inf.*	*Iteration*	*Feature*	*Cond. Tr. Inf.*	*Iteration*	*Feature*	*Cond. Tr. Inf.*	*Iteration*	*Feature*	*Cond. Tr. Inf.*
29	1	Sibil.	0.441	1	Sibil.	0.924	1	Dur.	0.392	1	Dur.	0.787
	2	Voice	0.331	2	Dur.	0.325	2	Sibil.	0.303	2	Sibil.	0.456
	3	Hi./An.	0.291	3	Hi./An.	0.453	3	Voice	0.339	3	Voice	0.597
	4	Cont.	0.123	4	Voice	0.598	4	High	0.104	4	Hi./An.	0.358
	5	Place	0.020	5	Cont.	0.198	5	Back	0.019	5	Cont.	0.146
				6	Place	0.094	6	Fric.	0.066	6	Nasal	0.116
							7	Anter.	0.063			
Total			1.206			2.591			1.286			2.461