INFORMATION TO USERS

This material was produced from a microfilm copy of the original document. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the original submitted.

The following explanation of techniques is provided to help you understand markings or patterns which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting thru an image and duplicating adjacent pages to insure you complete continuity.
2. When an image on the film is obliterated with a large round black mark, it is an indication that the photographer suspected that the copy may have moved during exposure and thus cause a blurred image. You will find a good image of the page in the adjacent frame.
3. When a map, drawing or chart, etc., was part of the material being photographed the photographer followed a definite method in "sectioning" the material. It is customary to begin photoing at the upper left hand corner of a large sheet and to continue photoing from left to right in equal sections with a small overlap. If necessary, sectioning is continued again — beginning below the first row and continuing on until complete.
4. The majority of users indicate that the textual content is of greatest value, however, a somewhat higher quality reproduction could be made from "photographs" if essential to the understanding of the dissertation. Silver prints of "photographs" may be ordered at additional charge by writing the Order Department, giving the catalog number, title, author and specific pages you wish reproduced.
5. PLEASE NOTE: Some pages may have indistinct print. Filmed as received.

Xerox University Microfilms
300 North Zeeb Road
Ann Arbor, Michigan 48106

73-22,740

MITCHELL, Patricia Dooley, 1943-
A TEST OF DIFFERENTIATION OF PHONEMIC FEATURE CONTRASTS.

The City University of New York, Ph.D., 1973
Speech

University Microfilms, A XEROX Company, Ann Arbor, Michigan

A TEST OF DIFFERENTIATION OF PHONEMIC FEATURE CONTRASTS

by

PATRICIA DOOLEY MITCHELL

A dissertation submitted to the
Graduate Faculty in Speech in partial
fulfillment of the requirements for the
degree of Doctor of Philosophy,
The City University of New York.

1973

This manuscript has been read and accepted for the Graduate Faculty in Speech in satisfaction of the dissertation requirements for the degree of Doctor of Philosophy.

May 21, 1973
date

Katherine S Harris
Chairman of Examining Committee

May 22, 1973
date

Norma S. Rees
Executive Officer

Katherine S. Harris
James Abel
Philip Hood
Supervisory Committee

The City University of New York

ACKNOWLEDGMENTS

The completion of the work described in this dissertation would not have been possible without the help of many generous persons.

First, I would like to thank my adviser, Professor Katherine S. Harris, not only for the ideas contributed to this study and her encouragement during its preparation, but also for introducing me to the field of speech science.

I also wish to thank the members of my dissertation committee: Professor James Abel, for his many suggestions and careful editing of this manuscript; and Professor Philip Hood, for his valuable advice.

I am also deeply indebted to: Professor Jean Louvrinic of the Temple University Medical School Audiology Clinic, for providing the clinical data and subjects; to Professor Thomas Gay of the University of Connecticut, for recording the test lists; to Professors Harry Levitt, Louis Gerstman, and Moe Bergman of the City University of New York, Professor Arthur Abramson of the University of Connecticut, Dr. Peter Bricker of Bell Laboratories, and Dr. Franklin S. Cooper and the staff of Haskins Laboratories, for their valuable advice and assistance; and to former and present fellow students Lorraine Russell, Dr. Fredericka Berti, Dr. Ernest Zelnick, and Professor Lawrence Deutsch, who have aided me in this study.

Thanks are also due to the students and faculty at the City University of New York, who served as the subjects in the listening experiments.

Finally, I wish to thank my husband, Bob, and my family for their understanding and assistance during the preparation of this manuscript.

Patricia Dooley Mitchell
May, 1973.

TABLE OF CONTENTS

LIST OF TABLES

Table Page

LIST OF ILLUSTRATIONS

CHAPTER I

INTRODUCTION

The most commonly used index of the ability to differentiate the sounds of speech is the articulation score. The percentage of the words or syllables spoken by a talker and identified correctly by a listener is calculated for a given set of listening conditions. The score provides a quantitative measure of speech intelligibility, whether the need is to assess the ability of a hard-of-hearing individual to understand speech, or to evaluate the efficiency of communications systems in the transmission of speech.

In recent years, several investgators have expressed a desire to gain more from such tests than a maximum score. These same articulation tests and formats, though not specifically designed for this purpose, have been used to study the errors occurring on particular speech sounds.

It is the purpose of the present study to develop an articulation test with speech materials and format that are appropriate for the analysis of the confusions among the sounds of speech.

The test used in this study is of the multiple-choice type. One of the 200 words spoken is "pat." The listener must identify what he has heard by selecting one of the following five words: "pat, bat, mat, fat, cat." Each of the foils differs from the stimulus by only one significant speech sound, known as a phoneme. In turn, each of the phonemes differs from the stimulus phoneme by only one phonemic feature: voicing, nasality, manner of articulation, and place of articulation, respectively. For the present discussion, any pair of phonemes differing from each other by only one feature is considered a minimal-feature contrast. This format allows tabulation of errors, not only by phoneme confusions, but also by feature confusions.

Problems in Articulation Testing

The discrimination of complex speech signals is undoubtedly the most important of the several different aspects of man's hearing. However, the use of speech as test material is complicated by many factors. For example, speech sounds have components of varying frequencies and they are constantly changing, or, even with modern electroacoustic measuring devices, it is difficult to quantify the speech signal's power at any given moment. One must generally choose some referent, such as the loudest portion of the word, whose level can be measured, and allow the rest of the signal to maintain its natural, but unspecified, relationship to this referent. One must also consider differences of speakers' voices, dialects, word familiarity, and room acoustics, as well as other variables.

Despite the inherent difficulties with speech materials, they provide the only valid test of speech discrimination. One cannot predict this ability from simpler measures such as pure-tone thresholds. For certain hearing disorders, especially those involving the central nervous system, pure-tone thresholds may be normal, while more sophisticated functions, such as speech discrimination, may be seriously impaired. Thus, articulation tests have been and will continue to be used in the evaluation of communication systems and in the clinical diagnosis and rehabilitation of auditory disorders.

The ease with which one can discriminate speech sounds is highly dependent on the nature of the materials tested. Sentences, for instance, are easier to understand than isolated words because of the contextual cues they contain. For much the same reason, disyllabic words are easier than monosyllabic words, which are in turn easier than nonsense syllables. Monosyllables depend more upon the differentiation of individual phonemes, since they have fewer contextual cues, while nonsense syllables depend even more on perception of the individual phonemes for their intelligibility, since meaningful phoneme sequences are not involved.

Articulation test experiments were first conducted by the telephone industry. Investigators calculated the percentage of nonsense consonant-vowel-consonant (CVC) syllables in which all three phonemes were heard correctly. Nonsense syllables could be used to advantage in these experiments since trained listening crews were employed.

For general use however, articulation lists of nonsense syllables are inappropriate because the responses will be biased by the relative familiarity of certain syllables unless the listeners are taught the entire list beforehand. Therefore, lists of actual English words were developed for wartime research on communication systems and for clinical use.

The best known of these lists was compiled on the principle of phonemic balance, which requires that the phonemes comprising the list have the same proportional distribution as they have in the language as a whole. These tests, referred to as PB lists, are those most commonly used for diagnostic and experimental purposes. The maximum score obtained on a PB list when presented at a comfortable listening level tells the clinician how well the patient can discriminate speech sounds when intensity is not the major factor. These scores are useful as an aid to differential diagnosis, since certain types of pathologies affect discrimination more severely than others, and as a guide to the selection of a hearing aid or other rehabilitational procedure.

For these same purposes, it would seem that information more detailed than a maximum score should be obtained. Further, it would be useful to know which phonemes or classes of phonemes contribute the greatest number of errors and which phonemic feature or features of those phonemes are confused. This kind of phonemic analysis might be far more diagnostically informative than equally time-consuming tests which yield only percentage scores. Recent research suggests that the errors made on articulation tests are not random, and therefore, not equal. Yet all errors on PB lists are treated as equal, being scored correct or incorrect. The omission of a final consonant, the confusion of acoustically similar consonants, or the substitution of an entirely different word contributes equally to the score.

The first attempts at error tabulations were analyses of errors on PB lists. However, these lists are inappropriate for analytic uses, chiefly because they are administered as open-response message sets; that is, for each word presented, the listener can choose any word in his vocabulary as his response. Therefore, the response is biased by the relative familiarity of words; one is less likely to choose an unfamiliar word. In addition, the examiner has no control over the phoneme being tested, because the listener may confuse any one of the component phonemes and respond with an entirely different word that makes sense of the confusion. Any analysis of the substituted word is also biased by the availability of words in the language which contain the phoneme heard in error. Therefore, the examiner has no control over the phonemes or words that can be substituted.

Until recently, most of these basic problems with PB lists have been ignored, while much research has been done on such questions as: the necessity of phonemic or phonetic balance, the achievement of better phonemic balance, the shortening of the lists without destroying the phonemic balance, and the effects of various recording techniques on the discrimination scores.

Within the last few years, several new speech discrimination tests have been developed for both clinical and experimental uses. The most important innovation has been the use of the closed-response set technique, in which the entire message set is defined for the listener. For each test item a limited number of choices is available, and the listener is forced to choose among them. Thus the effects of word familiarity are minimized without the necessity of extensive prior training, and the phonemes under test, as well as their alternatives, are controlled by the examiner.

For most of these closed-response tests, the criterion of phonemic balance has been abandoned. The emphasis has been shifted to the development of multiple-choice items for use in the tabulation of the phonemes heard incorrectly, and in some cases, an analysis of the substitutions.

Various models have been proposed for the description of phonemes. These are based on sets of features or dimensions, each of which varies independently in its value for a given phoneme. These features may relate to articulatory (productive), acoustic, or auditory (perceptual) aspects. The exact relationships between any two of these aspects has not been completely defined. One set of features based on productive aspects includes voicing, nasality, place of articulation, duration, and affrication. Values, such as "present" or "absent," or "front" or "back" assigned for each feature to each phoneme, yield a description for each phoneme which is unique but incomplete. This phonemic model, as well as others, will be described in greater detail later.

Interest in the place of phonemic features in the design of new articulation tests results from the greater success in the finding of more regular patterns of errors when phonemes are grouped and analyzed by their features than when phonemes are considered individually. All attempts thus far, whether clinical or experimental, to predict the particular phonemes which will be confused have met with little success. However, the results of several investigations over the past two decades have led to the conclusion that the perception of one phonemic feature is relatively independent of the perception of the others.

In addition, under known conditions of distortion of the speech signal, the errors are somewhat predictable. What is meant by "predictability" in such analyses is that the substitute phoneme will have one or more features in common with the correct phoneme. For example: /p/ is more likely to be confused with /k/, a one-feature contrast, than with /g/, a two-feature contrast. These predictions are based on some particular model of phonemic features. The extent to which such predictions will be accurate depends on the validity of the model, that is, the degree to which each of the features described is representative of an actual channel in the perceptual process.

Existent closed-response tests allow feature analysis, but only after a certain degree of subjective interpretation by the examiner. This subjectivity results from the fact that the phonemes which contrast the members of the test sets differ from one another in more than one feature. For example, on one such test developed by Griffiths (1967), the first set is "bat, batch, bass, bash, badge," and the stimulus is "bat." These are alternatives which differ from the stimulus in place of articulation, manner of articulation, manner and place, and voicing and place, respectively. (The alternatives "batch" and "badge" also differ in manner of articulation if the affricates are not classed as stops.) There are no alternatives providing one-feature contrasts of nasality or voicing. On another test list the stimulus word is "badge" and the same set of alternatives is presented. In this case any error would include voicing, regardless of the other features involved.

In order to tabulate objectively the number of errors for any given feature, each contrast should represent an error in one feature only. In the use of the currently available tests, multi-feature errors must be interpreted as one-feature errors. If the alternative chosen was the only one differing in a given feature, it is assumed that it was chosen for lack of information about that feature only. Using the example given above, if "badge" were given in response to "bat," the error was considered by Griffiths to be a pure voicing error, regardless of the difference in place, and a possible difference in manner if affricates are not classed as stops. Such subjective interpretive procedures could be avoided if "bad," a pure voicing contrast, were offered as an alternative.

One could, however, develop a test which provides all possible minimal-feature contrasts for any given stimulus, except that the complete set of entirely minimal contrasts could not be generated for certain classes of phonemes. Since there are no voiceless nasals or semivowels, any voicing contrast for these sounds would be inherently multidimensional. However, the need for interpretation by the examiner would be minimized since these multi-feature contrasts would be assigned as particular minimal-feature errors in advance.

Studebaker (1968) found with another closed-response test that certain feature contrasts remain intact for most hard-of-hearing listeners. Only those with severe speech discrimination problems confused the voicing or nasality features. Therefore, Studebaker developed another test limited to contrasts of the place of articulation feature for use with most hard-of-hearing listeners. A test of easily differentiated voicing and nasality contrasts was reserved for those cases with severe discrimination problems. It would seem, however, that one test could be devised which would be appropriate for all types of cases and one which would indeed be necessary for any cross-pathology studies or differential diagnoses.

Studebaker's test also allows omission of the phoneme as a response. If omissions are not allowed, greater significance for the substitutions may be found. Hence, new tests should not permit omissions, as they do not reveal what a patient hears for a phoneme, and in one study, it was found that neither omissions nor additions differentiated among pathologic groups. This seems to have come about because a patient heard some cue, such as the transition indicating manner, but did not have enough information to guess without alternatives. Omission of a phoneme as a response does not necessarily mean that the listener heard nothing.

Furthermore, there does not appear to be any single test for the analytic study of errors on word-initial and final consonants, as well as on vowels. Most closed-response tests involve consonant discrimination only. Yet Schultz (1964) found "...a sufficient number of [vowel] error differences among otopathologies to merit serious consideration of a test requiring vowel discrimination in addition to present instruments." A few investigators have developed two multiple-choice tests, one for consonants and one for vowels, each with different types of materials and formats.

New Test Criteria

In light of these developments, it appears that for the analysis of phoneme errors there is needed an experimental and clinical test which meets the following criteria:

Vocabulary.- The test vocabulary should consist of actual English words, since nonsense syllables are inappropriate for use with untrained subjects. The sample of test items should include both vowels and word-initial and word-terminal consonants. In addition, the environmental influence of phonemes adjacent to the test phoneme should be considered.

Format.- The format should be of the closed-response type in order to account for varying word familiarity. More importantly, this technique allows testing of one phoneme at a time.

The administration of the test and the task of the listener should be consistent throughout; that is, the format and the number of alternatives should be the same for all consonants and vowels. In this way, the probability of error is the same for all classes of phonemes. Each test item should have its own set of alternatives, selected in such a way that all possible one-feature contrasts are given, since these errors have been shown to be more likely than multi-feature errors. Omissions or additions should not be allowed, since they yield little or no information about the distortion of the phonemes presented.

Scoring.- The choice of any alternative word in a set should be established in advance as constituting a given type of error, thus minimizing subjective interpretation of errors by examiners.

Applicability.- The same test should be appropriate for subjects with discrimination difficulties ranging from mild to severe. Since certain one-feature errors appear to be more likely than others, the test should provide for both relatively easy and relatively difficult discriminations, such as voicing and place of articulation. Thus the same test could be used for analyses of discrimination errors resulting from all types of speech distortion, whether related to electroacoustic equipment, filtering, or hearing losses of varying types and degrees.

It is the purpose of the present investigation to develop, for use in the analyses of speech sound confusions, a test which more nearly meets the above criteria than does any existent test.

Such a test may provide a quantitative method of describing the predictability of feature errors in terms of the particular features which are lost under given conditions. For example, the elimination of certain portions of the spectrum may be associated with a predominance of errors in one particular feature. If such a model can be established for normal listeners when certain portions of the frequency spectrum are filtered out, cases of various otopathologies can be compared to the norm in order to investigate the diagnostic significance of the types of errors. Information obtained by this test might also be useful in summarizing the rehabilitative needs of an individual with difficulties in speech discrimination.

CHAPTER II

ARTICULATION TESTING

The Need for Tests of Speech Discrimination

Sensitivity for pure tones was the earliest aspect of hearing tested in a standardized manner. The intensity levels at which normal hearers should be able to detect the presence of a tone of a given frequency were established by obtaining these thresholds for large samples of listeners (Beasley, 1938; Steinberg et al., 1940). The resulting average levels were then used in the calibrations of the electric audiometer. Any deviation from these "normal" thresholds indicated a hearing loss which could be specified in decibels (dB).

However, the criterion for determining the degree to which a hearing loss constituted an impairment or a handicap was the effect of the loss on the ability to understand speech. Hence, the series of several discreet frequency thresholds shown on the audiogram had to be converted into a single measure which would indicate how well the individual could understand speech under normal listening conditions.

Previous articulation testing under controlled conditions of intensity and frequency distortion had shown that the frequencies yielding the greatest contribution to speech intelligibility were located between 250 and 3500 Hz (Fletcher, 1929; Fletcher and Steinberg, 1929; and Pollack, 1948). With these important speech frequencies established, it was decided to average the thresholds for the frequencies 500, 1000, and 2000 Hz as an index of the hearing level for speech (Carhart, 1946b). Various methods of weighting the more important frequencies in determining the average were devised (Fletcher, 1950).

Speech audiometry was developed for use in school screening tests and for military use. Recorded or live samples of spoken digits were attenuated until the listener could identify the stimuli 50 per cent of the time. This level, known as the speech reception threshold (SRT), yielded an estimate of the hearing level for speech stated in terms of a single number. Standardized methods of obtaining the speech reception threshold showed that it was a rather stable measure. Young and Gibbons (1962) found a correlation of r=.90 between the threshold at 500 Hz and the SRT. Thus, the SRT can serve as a confirmation of the audiogram and as an aid to differential diagnosis. Infact, marked discrepancy between the pure-tone average and the SRT is in itself diagnostic of psychological or central disorders.

However, the ability to identify speech stimuli at threshold levels is not an adequate measure of the individual's ability to understand speech in everyday situations. A more valid question about normal communication is, for a given intensity, how intelligible is a sample of speech, or how well can a person differentiate the sounds of speech. As Kryter _et al_. (1962) point out, at threshold levels, the frequency components of speech about 1800 Hz are near the threshold of sensitivity for even normal ears, and thus, the threshold performance of impaired ears will not **differ** significantly from that of the norm. However, the performance of impaired ears on a test of speech discrimination may be significantly different. This cannot be predicted from the audiogram or the SRT.

The frequencies whose average was the best predictor of the speech discrimination score were found to be different by each of several investigators. Mullins and Bangs (1957) found that the pure-tone average of 2000 and 3000 Hz predicted the PB score better than any other combination. Kryter _et_ _al_. found that the average of 2000, 3000, and 4000 Hz was best, with a correlation of $r=.81$ ($r^2=.66$). Young and Gibbons found that the $r=.65$ for the threshold at 2000 Hz showed the highest correlation with the PB score. They concluded that the index of determination ($r^2=.42$), which this yielded, was hardly large enough to warrant the use of a pure-tone threshold as a predictor of discrimination scores. Even when all subjects with an 88% discrimination score were grouped, there was a wide variety of pure-tone and speech reception thresholds. Pickett _et_ _al_. (1972) found a correlation of $r=.66$ between the average of thresholds at 1000 and 2000 Hz and the scores on their modification of the Fairbanks Rhyme Test.

Thus, it appears that the only valid test of speech discrimination is the presentation of speech materials at supra-threshold levels.

The Development and Use of Early Speech Discrimination Measures

The development of speech materials for determining discrimination scores began with lists of nonsense (CVC) syllables used at Bell Telephone Laboratories (Fletcher and Steinberg). Fletcher (1929) also developed lists of sentences in which repetition of only certain key words was scored. Other investigators constructed lists of real words for clinical uses, e.g., Fry and Kerridge (1939).

World War II further stimulated interest in articulation testing as a means of evaluating communication systems. Lists of CVC words were developed for this purpose at the Harvard Psycho-Acoustic Laboratory. The basic criterion in the development of these lists was that of "phonetic balance" (Egan, 1948). In order to satisfy this criterion, each list had to contain fifty words. These lists are therefore generally referred to as the "PB-50" lists.

Due to the extensiveness of the vocabulary and the low interlist comparability on the Harvard lists, modified lists were developed by Hirsh and his colleagues at the Central Institute for the Deaf (CID), (Hirsh *et al.*, 1952).

In the interest of achieving greater phonetic balance, Lehiste and Peterson (1959) devised their CNC (consonant-nucleus-consonant) Lists. These were based on considerations not only of the distribution of phonemes but also of the distribution of consonant-vowel combinations in the language. Lehiste and Peterson claimed that their lists were more "phonetically" balanced, while the Harvard and CID lists were actually only "phonemically" balanced. These CNC lists were later revised to include more familiar words (Peterson and Lehiste, 1962).

All of the tests described above were designed primarily for use in obtaining a maximum articulation score, that is, the percentage of syllables or words heard correctly.

These tests have been used also in measurement of the relative contributions of different frequencies to speech intelligibility (Fletcher; Fletcher and Steinberg; and Pollack, 1948). These investigations resulted in the saving of half the bandwidth for transmission of speech signals via radio or telephone. Although the speech spectrum extends from below 100 Hz to about 7000 Hz, only those frequencies between 250 Hz and 3500 Hz need to be transmitted to achieve intelligible communication (Licklider and Miller, 1951). Such findings imply that these end portions of the speech spectrum are redundant. The normal listener receives the necessary information without these extra cues.

The distribution of this information within the frequency spectrum was studied by French and Steinberg (1947). The results of their articulation tests with CVC nonsense syllables revealed that the same scores were obtained when frequencies below 1900Hz were passed (low-pass filtering) as when frequencies above 1900 Hz were passed (high-pass filtering). This cross-over point of 1900 Hz represents the frequency above and below which equal amounts of information are transmitted for the type of material tested. Cross-over points almost an octave lower were found when speech was masked by bands of noise (Pollack, 1948; Kryter et al., 1962; and Webster and Klumpp, 1963). An even lower cross-over point of 725 Hz was found for synthetic sentences (Speaks, 1967). The data of French and Steinberg indicated that the most important frequencies for over-all intelligibility of monosyllables were between 1500 Hz and 2500 Hz, while Speaks' data showed that for sentences the contribution of frequencies over 1000 Hz was minimal. Therefore the critical portion of the spectrum is far lower for sentences, which carry contextual cues, than for monosyllables, which carry little, if any, contextual information.

These studies show that although certain portions of the spectrum are more important for overall intelligibility, the determination of these portions are dependent on the types of material used. Thus, one can predict the overall intelligibility of specified materials transmitted via communication systems with known frequency response.

However, one cannot similarly predict the intelligibility scores obtained for hard-of-hearing subjects whose frequency response, in terms of the audiogram, is also known. This may result because not all speech cues are frequency dependent. Other cues operate in speech perception, such as duration or the rate of frequency changes in the transitions from one phoneme to another. In addition, although comparisons between frequency responses of electroacoustic filters and hard-of-hearing subjects are a convenient starting point for experimental studies, one cannot think of the hard-of-hearing listener as if he were hearing speech with certain frequencies filtered out. Other types of distortion in the pathologic ear may cause discrimination to be poorer than would be expected from filtering experiments, while loudness recruitment or experience with a hearing disability may cause discrimination to be better than expected.

Rhodes (1966) studied the responses by normal and hard-of-hearing listeners to filtered words. He presented CNC words through low-pass filters set at 1000 Hz to a group of hypacusics with losses above 1000 Hz and to a group of normal controls. A highly significant difference in favor of the hypacusics was found. With no overlap in the individual scores, the mean score for the hypacusic group was 58.8%, while the mean score for the normals was 30.8%. No significant difference was found between a hypacusic group with losses above 2000 Hz and their normal controls when the low-pass filter was set at 2000 Hz. Their scores were 93.4% and 93.6%, respectively. It seems evident that individuals with certain high-frequency hearing losses utilize cues not recognized by normal listeners.

It must be concluded, therefore, that the only means of estimating how well a given hard-of-hearing listener will discriminate speech is by testing this function directly.

The Perception of the Individual Speech Sounds

The effects of frequency distortion on the individual speech sounds was studied by Fletcher (1929, 1953). The results of this analysis showed that some of the sounds have characteristics localized in a limited frequency region, while other sounds seem to have characteristics extending throughout the entire frequency range. Vowels and diphthongs have sufficient distinguishing characteristics in either half of the frequency range. The vowel /i/, for example, was recognized correctly 98% of the time when either the frequencies above 1700 Hz or below 1700 Hz were passed. The consonants, however, were far more adversely affected by elimination of the high frequencies. The /s/ sound was only slightly affected by elimination of frequencies below 1500 Hz, but its correct recognition was reduced to 40% when frequencies above 3000 Hz were eliminated.

As Fletcher's data suggest, there is an association between a given phoneme and one or more portions of the frequency spectrum. The locations of these energy concentrations have been determined in a rather straightforward way by spectrographic analysis. Peterson and Barney (1952) measured the frequency of these energy concentrations or resonance bands (known as formants) for ten vowels, averaging those stimuli classified correctly by all 70 listeners. A rise in the first formant frequency is associated with an increase in the openness of the vocal tract during the production of a vowel. Thus, low-open vowels have high first formants, while high-close vowels have low first formants. A rise in the second formant frequency is associated with a place of articulation that is further forward in the vocal tract.

Thus, front vowels have high second formants, and back vowels have low second formants. In the Peterson and Barney data there is a good deal of overlap and scatter in the distribution of the first and second formants. It is most extensive for /ʊ/ and /ɝ/. However, /ɝ/ is probably easily differentiated from the other vowels by its low third formant.

In another study, Peterson (1952) found that the third formant contributes somewhat to the intelligibility of all vowels, but is especially important for the separation of the front vowels. It thus appears that more than just the first two formants are needed to specify the acoustic correlates of perceptual distinctions among vowels.

The possibility of a different perceptual hierarchy among the formants of different vowels has been suggested by Chiba and Kajiyama (1938). They reported that when five vowels were presented through various band-pass filters, all vowels were perceived as one particular vowel for each of the pass bands. The band coincided with one of the formant regions for that vowel. This notion of a principal formant for each vowel was verified by Castle (1964). He found that correct recognition of /i/ was most apt to occur when a band of frequencies surrounding its third formant was presented. The resulting hierarchy of F3-F2-F1 was also found for /æ/. For the vowel /ɑ/, the order of importance was F2-F1-F3, and for /u/ it was F1-F2-F3.

The degree to which formant frequencies may vary for different utterances of the same vowel has been shown to be quite large (Chiba and Kajiyama; Potter and Steinberg, 1950; and Peterson and Barney, 1952). It appears that there is an allowable range of frequencies for the formants of each vowel. A summary of the spoken vowel response data for thirteen band-pass filter settings reported by Castle (1964) supports this hypothesis. With the single exception of /ɛ/, each of the vowels was associated with each of several consecutive frequency bands. For /i/, there were two sets of consecutive bands, one including the two uppermost (3150 and 2500 Hz) and one including the four lowermost (400, 315, 250, and 200 Hz). Disregarding the lower set for /i/, the sets of bands become progressively lower for /i/, /æ/, /ɑ/, /ɔ/, and /u/. Within each set there was one band which appeared to be most closely associated with the perception of a given vowel.

These results obtained by filtering spoken vowels are in essential agreement with results obtained by synthesizing one-formant vowels. Delattre et al. (1952) reported that when low formant frequencies were presented, back vowels with corresponding first formants were heard. When high formant frequencies were presented, front vowels were assigned according to their second formants.

In addition to formant frequency, there are other cues which appear to operate in vowel perception. Among these is intrinsic duration. Though the durations of vowels vary widely with the speaker, context, and tempo, there appears to be an intrinsic duration associated with each vowel. Measurements of the vowel lengths for speakers from three different American dialectal areas showed in all cases a "short" series, including: /ɛ, ʊ, ɪ, ʌ/; and a "long" series including: /ɔ, i, eɪ, u, æ, ɝ, oʊ, aʊ, a, aɪ/ (Locke and Heffner, 1940; and Lehmann and Heffner, 1940).

Peterson and Lehiste (1960) define the intrinsic duration of syllable nuclei (vowels and diphthongs) as ". . . the average duration of the syllable nucleus measured from minimal pairs differing in the voicing of the final consonant. . . ." (p.702). In this manner, they established groups of intrinsically short and intrinsically long syllable nuclei, and these groups were identical to those established by Locke and Heffner and Lehmann and Heffner.

As Mattingly (1968) suggests, the intrinsic durations of vowels may be a subsidiary cue to their perception, since vowels with similar spectra differ in length, as in the cases of /i/-/ɪ/ and /u/-/ʊ/.

Miller (1956) analyzed the confusions of a series of spoken vowels presented through a 670 Hz low-pass filter. He assumed that only the first formant information would be transmitted. The listeners tended to confuse /ɪ, ɛ, ʌ, ʊ/ with each other, and /i, æ, ɑ, ɔ, u/ with each other.

The confusions of vowels heard in various bands of noise (Pickett, 1957) verify the role of duration in vowel perception. The use of either high-frequency or flat noise caused confusions of the second formants, while the use of low-frequency noise caused confusions among the first formants. However, the particular confusions suggested some joint operation of duration cues and formant frequency.

The relative intensity of the formants may also be a cue to the perception of vowels. Delattre *et al.* (1952) varied the intensities of the formants of synthetically-produced vowels. When the first two formants were far apart, as in front vowels, the reduction of the intensity of the first formant resulted in a loss of vowel-like quality, but did not cause a shift in vowel color. In cases where the formants were close together, as in back vowels, reduction of intensity of the first formant resulted in a shift toward an adjacent vowel. The authors suggested that those back vowels with formants close enough for perceptual averaging are, in effect, one-formant vowels. Therefore the reduction of the intensity of one of the formants decreased its contribution to the average, causing a shift to another vowel. When the second formant was reduced in intensity, the result for all vowels was a shift to back vowels with similar first formants.

In the case of a diphthong, the absolute formant frequency seems to be even less important than for pure vowels. Conventionally, diphthongs have been described as glides from one vowel position to another. This implies that the important acoustic cues would be the formants of the initial vowel position and the terminal vowel position. The results of a study of synthetically-produced diphthongs (Gay, 1967) suggest that the rate of formant change is the important acoustic cue. A later study confirmed this (Gay, 1968). When diphthongs were spoken at slow, moderate, and fast rates, the rate of formant change was found to be constant. For fast speaking rates, the diphthong ceased before reaching the terminal vowel target.

Although context influences the formant frequencies and durations of vowels, one can for descriptive purposes, consider some steady portion of the vowel. However, for the consonants, one must consider not only the frequency spectrum of the consonantal-noise segment, but also the formant transitions to or from the adjacent vowels.

For consonants, the first formant frequency and the rate of its transitional change from vowel to consonant, or consonant to vowel, is a cue to the manner of articulation. For example, a transition from the lowest possible first formant frequency to the steady-state first formant of the following vowel is a marker for the class of voiced stops. (For voiceless stops the first formant transition is not present.) The first formant begins at the lowest point for the stops, because they involve a complete obstruction in the vocal tract. For looser constrictions, the first formant begins at a somewhat higher frequency, as for fricatives (Delattre *et al*., 1955). The duration of transitions, that is, the rate of formant change, is also a cue to manner. Liberman *et al* (1954) found that increasing the transition duration caused some stops to be transformed into glide consonants, and with further increases, into glide vowels.

The cue to place of articulation of consonants is the transition of the second formant. Liberman _et al_. (1954) and Delattre _et al_.(1955) showed that the direction and degree of second formant transitions could serve as cues for the perceived distinctions among the stop consonants. The transition "locus," defined as the frequency to which the formant transition appears to point, is the same for all consonants with the same place of production. For example, the best second formant locus for synthetic /p/, /b/, and /m/ is around 720 Hz, while for /t/, /d/, and /n/, it is around 1800 Hz. The second formant locus for the velar consonants, /k/, /g/, and /ŋ/, varies according to the place of articulation of the adjacent vowel.

The spectrum of the noise burst which accompanies the articulatory explosion may also serve as a cue to the place of articulation for consonants. Liberman _et al_. (1952) found that bursts above 3000 Hz were generally perceived as /t/. Below that position, the perception of the burst was dependent on its relationship to the adjacent vowel.

A study of the effects of consonant noise versus vowel transition on the perception of fricatives was reported by Harris (1958). The stimuli were constructed by combining the noise from one spoken fricative-vowel syllable with the vowel portion of another (including the transition), and the listeners were asked to identify the resulting sounds. The results suggested that /s/ and /ʃ/ were dependent on the noise, while /f/ and /θ/ were differentiated by the second formant transition into the adjacent vowel. A later study of synthetic syllables by Heinz and Stevens (1961) led to similar conclusions.

Experimental evidence (Lisker and Abramson, 1964 and 1967) has led to the conclusion that the distinction between voiced and voiceless consonants appears to be a matter of the timing of changes in the glottal opening, relative to the vocal tract articulation. The acoustic correlate of this articulatory relationship is the "voice onset time," defined as the time interval between the noise segment that marks the release of the consonant and the onset of periodicity which reflects laryngeal vibrations. The English voiced consonants /b/, /d/, and /g/ would be characterized by an onset of voicing before or immediately after the consonant release. The voiceless consonants /p/, /t/, and /k/ would exhibit a delay of at least 30 milliseconds between the consonant release and the onset of voicing. An additional cue to the presence of voicing of post-vocalic consonants may be the duration of the preceding vowel. Any vowel followed by a voiced consonant will be longer than the same vowel followed by a voiceless consonant. Peterson and Lehiste (1960) found, as others had previously, that the voiced-voiceless contrast influence was clearly evident for both intrinsically long and intrinsically short vowels.

In summary, the place-of-articulation cues of all phonemes are frequency dependent. In the case of consonants, some appear to be more dependent on the frequency of the vowel transition, while others appear to depend on the frequency of the noise which accompanies the consonant release. Manner of articulation is also dependent on transition frequencies, as well as the rate of transition change. The cue to the perception of the voicing feature for consonants appears to be dependent on timing.

The description given here of the acoustic cues which give rise to the perception of the phonemes is not intended to be complete. The cues discussed are the acoustic characteristics that generally have been shown to be most critical to phoneme separation. If such a cue separates one group of phonemes from another, it can be considered to be the acoustic correlate of a phoneme feature. Such features may include: voicing, place of articulation, manner of articulation, duration, nasality, and others which will be discussed in the next section of this chapter.

The Role of Distinctive Features in Speech Analysis

In order to develop a parsimonious description of the characteristics of the phonemes it is useful to group them according to the phonemic features which distinguish them. These features are usually defined according to the articulatory processes which correlate, in theory at least, with the acoustic characteristics.

According to the distinctive feature concept outlined by Jakobson, Fant, and Halle (1963), when two phonemes differ from one another by one distinction which cannot be further resolved, the distinction is called "minimal." Wider differences are termed "duple," triple," etc., according to the number of minimal distinctions involved. The perception of any phoneme involves the resolution of a set of features, which, according to Jakobson _et al_., is always done by a process of binary decisions. Since binary theory imposes the restriction that no feature have more than two values, "present" or "absent," a large number of features is required to complete the necessary description. Table 1 presents the distinctive feature values for the consonants in the Jakobson _et al_. system.

This feature system has been revised by Chomsky and Halle (1968). The newer system, outlined in Table 2, is more complex because description of other than English phonemes requires a greater number of features. The terminology for some of the features has been changed. For example, "high, low, back" correspond to the former "diffuse, compact, grave." The authors intended the features to be abstract classificatory devices, and, on this level, the features are binary. However, on the phonological level, in both acoustic and articulatory terms, Chomsky and Halle admit that phonetic features are generally multivalued (1968).

To test the validity of distinctive feature theory, Wickelgren analyzed the confusions in short-term-memory for English vowels (1965a) and for consonants (1965b). The results of both studies led to the conclusion

TABLE 1

THE JAKOBSON-FANT-HALLE DISTINCTIVE FEATURE SYSTEM FOR CONSONANTS[a]

Consonant	+ Vocalic - Non-Voc.	Consonantal Non-Conson.	Compact Diffuse	Grave Acute	Nasal Oral	Tense Lax	Continuant Interrupted	Strident + Mellow -
/p/	-	+	-	+	-	+	-	-
/b/	-	+	-	+	-	-	-	-
/m/	-	+	-	+	+	-	-	-
/t/	-	+	-	-	-	+	-	-
/d/	-	+	-	-	-	-	-	-
/n/	-	+	-	-	+	-	-	-
/tʃ/	-	+	+	-	-	+	-	+
/dʒ/	-	+	+	-	-	-	-	+
/k/	-	+	+	+	-	+	-	-
/g/	-	+	+	+	-	-	-	-
/f/	-	+	-	+	-	+	+	+
/v/	-	+	-	+	-	-	+	+
/θ/	-	+	-	-	-	+	+	-
/ð/	-	+	-	-	-	-	+	-
/s/	-	+	-	-	-	+	+	+
/z/	-	+	-	-	-	-	+	+
/ʃ/	-	+	+	-	-	+	+	+
/ʒ/	-	+	+	-	-	-	+	+
/w/	-	-	-	+	-	-	+	-
/r/	+	+	+	-	-	-	+	-
/l/	+	+	-	-	-	-	+	-
/j/	-	-	-	-	-	-	+	-
/h/	-	-	+	+	-	+	+	-

[a] Based on Jakobson, Fant and Halle (1963), p. 43.

TABLE 2

THE CHOMSKY-HALLE (CH) DISTINCTIVE FEATURE SYSTEM FOR CONSONANTS[a]

Consonant[b]	Vocalic	Consonantal	High	Back	Low	Anterior	Coronal	round	voiced	Continuant	nasal	Strident
/p/	-	+	-	-	-	+	-		-	-	-	-
/b/	-	+	-	-	-	+	-		+	-	-	-
/f/	-	+	-	-	-	+	-		-	+	-	+
/v/	-	+	-	-	-	+	-		+	+	-	+
/m/	-	+	-	-	-	+	-		+	-	+	-
/r/	+	+	-	-	-	-	+		+	+	-	-
/l/	+	+	-	-	-	+	+		+	+	-	-
/t/	-	+	-	-	-	+	+		-	-	-	-
/d/	-	+	-	-	-	+	+		+	-	-	-
/θ/	-	+	-	-	-	+	+		-	+	-	-
/ð/	-	+	-	-	-	+	+		+	+	-	-
/n/	-	+	-	-	-	+	+		+	+	+	-
/s/	-	+	-	-	-	+	+		-	+	-	+
/z/	-	+	-	-	-	+	+		-	+	-	+
/tʃ/	-	+	+	-	-	-	+		-	-	-	+
/dʒ/	-	+	+	-	-	-	+		+	-	-	+
/ʃ/	-	+	+	-	-	-	+		-	+	-	+
/ʒ/	-	+	+	-	-	-	+		+	+	-	+
/k/	-	+	+	+	-	-	-	-	-	-	-	-
/g/	-	+	+	+	-	-	-	-	+	-	-	-
/ŋ/	-	+	+	+	-	-	-		+	-	+	-
/w/	-	-	+	+	-	-	-	+	+		-	
/j/	-	-	+	-	-	-	-	-	+		-	

[a] From Chomsky and Halle (1968), pp. 176-7.

[b] Phoneme symbols have been changed to agree with those used elsewhere in this paper.

that phonemes are coded in short-term memory, not as units, but as sets of distinctive features, each of which may be forgotten at least semi-independently. Some of the features of a phoneme can be recalled when others cannot, allowing predictions of substitutions which have some features in common with the phoneme presented.

Wickelgren compared several feature systems in terms of the accuracy with which each could predict the errors to be made. In his vowel study, there were no differences in the predictions made by a widely employed descriptive system which Wickelgren refers to as "conventional phonetic analysis" (CPA), and the Chomsky-Halle system (CH), since only six vowels were described.

Table 3 presents the descriptions of both CPA and CH systems for the six vowels in question. Place of articulation is handled in CPA by one feature with two values, front and back. Vowels are similarly divided in CH by the feature "grave," which can be present or absent. Openness of the vocal tract is handled in CPA by one feature with three values, narrow, medium, and wide. This dimension is handled in CH by two features, "diffuse" and "compact." For each of the six vowels, both systems predicted that the three vowels with one feature in common with the stimulus vowel will occur more frequently as response errors than the two vowels with no feature in common with the stimulus vowel. When the vowels were presented to listeners for recall, all such predictions were confirmed. Wickelgren also found that the

TABLE 3

CONVENTIONAL PHONETIC ANALYSIS (CPA) AND THE CHOMSKY-HALLE (CH) FEATURE SYSTEM[a]

	front	back	
+ diffuse	/ɪ/	/ʊ/	narrow
- diffuse - compact	/ɛ/	/ʌ/	medium
+ compact	/æ/	/ɑ/	wide
	- grave	+ grave	

[a] Reproduced from Wickelgren (1965a), p. 584.

correct value for openness was more likely to be forgotten than the correct value for place of articulation.

In his consonant study, Wickelgren compared three systems: Wickelgren (1965b), Miller and Nicely (1955), and Chomsky and Halle (1965). The Wickelgren (W) and Miller and Nicely (MN) systems, presented in Tables 4 and 5, respectively, are more similar to conventional phonetic description than is the Chomsky and Halle (CH) system.

Wickelgren used four features. Openness is the feature used to handle what is usually termed manner of articulation. In this system, affricates, stops, and nasals have the least degree of openness; most of the fricatives have a medium degree of openness; and the glides /w,r,j/, the lateral /l/, and the fricative /h/, have the greatest degree of openness. The place of articulation feature is handled by five values from front to back: (0)- /p,b,m,f, v,w/; (1)- /t,d,n,r,θ,ð/; (2)- /s,z,l/; (3)- /tʃ, dʒ,ʃ,ʒ,j/; (4)- /k,g,h/. Although more than five places can be identified, Wickelgren finds these sufficient to provide a unique description of the consonants. This system differs from conventional description in the assignments of phonemes to the (1) and (2) values. Generally, /s,z/ are grouped with /t,d,n,l/, while /θ,ð/ are placed in front of this group; /r/ is usually assigned a place farther back than any of these; and /h/ is not velar, as are /k,g/. However, the addition of another value for /h/ appears unnecessary since it also differs from /k,g/ by manner of articulation.

TABLE 4

THE WICKELGREN (W) FEATURE SYSTEM FOR CONSONANTS[a]

Consonants	Voicing	Nasality	Openness	Place
/p/	0	0	0	0
/b/	1	0	0	0
/m/	1	1	0	0
/t/	0	0	0	1
/d/	1	0	0	1
/n/	1	1	0	1
/tʃ/	0	0	0	3
/dʒ/	1	0	0	3
/k/	0	0	0	4
/g/	1	0	0	4
/f/	0	0	1	0
/v/	1	0	1	0
/θ/	0	0	1	1
/ð/	1	0	1	1
/s/	0	0	1	2
/z/	1	0	1	2
/ʃ/	0	0	1	3
/ʒ/	1	0	1	3
/w/	1	0	2	0
/r/	1	0	2	1
/l/	1	0	2	2
/j/	1	0	2	3
/h/	0	0	2	4

[a] Reproduced from Wickelgren (1965b), p. 390.

TABLE 5

THE MILLER AND NICELY (MN) FEATURE SYSTEM FOR CONSONANTS[a]

Consonant	Voicing	Nasality	Affrication	Duration	Place
/p/	0	0	0	0	0
/t/	0	0	0	0	1
/k/	0	0	0	0	2
/f/	0	0	1	0	0
/θ/	0	0	1	0	1
/s/	0	0	1	1	1
/ʃ/	0	0	1	1	2
/b/	1	0	0	0	0
/d/	1	0	0	0	1
/g/	1	0	0	0	2
/v/	1	0	1	0	0
/ð/	1	0	1	0	1
/z/	1	0	1	1	1
/ʒ/	1	0	1	1	2
/m/	1	1	0	0	0
/n/	1	1	0	0	1

[a] Reproduced from Miller and Nicely (1955), p. 347.

Miller and Nicely introduce the feature of duration, which they describe as an arbitrary term adopted to distinguish /s,z,ʃ,ʒ/ from the other consonants. MN use the feature "affrication" to distinguish fricatives from stops. Miller and Nicely define "affrication" as the acoustic turbulence resulting from air being forced through articulators which are close together. By contrast with W, the features affrication and duration cover manner of articulation. The MN place of articulation feature has three values: (0)- /p, b,f,v,m/; (1)- /θ,ð,n,t,d,s,z/; (2)- /k,g,ʃ,ʒ/. These three categories are considered sufficient to provide a unique description, since MN describes only sixteen consonants. Additional features or values would be required to handle the glides, the affricates, and /h/.

The voicing and nasality features are the same for both the W system and the MN system.

Since Wickelgren allows as many as five values for the place feature, and Miller and Nicely allow as many as three for the same feature, their descriptions require fewer features than does the Chomsky and Halle system, which requires that all features be binary.

The Chomsky-Halle feature system for consonants and vowels is presented in Tables 2 and 6, respectively. Actually, some of these features are further expanded into more features in their text. For example, Chomsky and Halle describe four binary features to handle the voicing distinction. These include: ± voicing, ± tense, ± sub-glottal pressure,

TABLE 6

THE CHOMSKY-HALLE DISTINCTIVE FEATURE SYSTEM FOR VOWELS[a]

Vowel[b]	Vocalic	High	Back	Low	Anterior	Coronal	Round	Tense
/i/	+	+	-	-	-	-	-	+
/ɪ/	+	+	-	-	-	-	-	-
/e/	+	-	-	-	-	-	-	+
/ɛ/	+	-	-	-	-	-	-	-
/æ/	+	-	-	+	-	-	-	-
/ɑ/	+	-	+	+	-	-	-	+
/ɔ/	+	-	+	+	-	-	+	-
/o/	+	-	+	-	-	-	+	+
/ʊ/	+	+	+	-	-	-	+	-
/u/	+	+	+	-	-	-	+	+
/ʌ/	+	-	+	-	-	-	-	-

[a] From Chomsky and Halle (1968), p. 176.

[b] Phoneme symbols have been changed to agree with those used elsewhere in this paper.

and ± glottal constriction. This expansion was necessary to support the binary framework when explaining the findings of other investigators, such as Lisker and Abramson (1964).

The only feature which appears to be the same for all three systems (CH, W, and MN) is the nasality feature.

The hypothesis tested by Wickelgren's analysis was that phonemes with a feature in common will be confused with each other more often than phonemes which differ in that feature. For example, there should be fewer voiced-voiceless confusions that would be expected from random data. While all three systems were significantly better than chance in predicting substitutions, Wickelgren concluded that short term memory uses articulatory or acoustic codes, such as those described by W or MN, and not the complicated abstract code described by CH.

Wickelgren's data were reanalyzed in terms of binary features by Klatt (1968). Using a similarity metric, Klatt tested the validity of the CH distinctive features. By Klatt's procedure, the dependency of one feature upon another (redundancy) could be measured and the existence of independent features could be determined. If the number of cross feature confusions in Wickelgren's data was significantly less than expected by chance, that feature was assumed to be present in the data. Those CH features which appeared to be present strongly and independently in the data were: long frication, sonorant, voiced, and continuant. Features that were absent or only weakly present included: consonantal, nasal, coronal, anterior, and strident. Klatt found no successful way of grouping consonants into binary place features. He did find, however, that a multivalued place feature was strongly evident. He concluded that places of articulation apparently do not group together into binary features at the perceptual level involved in short-term memory. This accounts for the poor performance of the CH system in the Wickelgren study.

Miller and Nicely's analysis of perceptual confusions suggested that a similar set of features is used for auditory perception as is used for coding in memory. Miller and Nicely had analyzed the confusions among sixteen consonants presented through various high-pass, low-pass, and band-pass filter settings at various signal to noise ratios. Each stimulus consisted of a consonant spoken initially before the vowel /ɑ/. Each consonant had a chance to be confused with every other consonant, and a count of such confusions was made. Having converted the consonant into a set of phonemic features (as described in Table 5), Miller and Nicely found that: "With noise or low-pass filtering the confusions fall into consistent patterns, but with high-pass filtering the errors are scattered quite randomly" (p.338). They found that when frequencies above 1000 Hz are filtered out, the information for the place and duration features is effectively lost, while the information for nasality, voicing, and affrication is preserved.

By this procedure, one can determine not only how much information is contained in a given portion of the spectrum, but also what kind of information. Miller and Nicely found that the cross-over points of equal contributions to intelligibility were different for the individual features: 450 Hz for nasality, 500 Hz for voicing, 750 Hz for affrication, 1900 Hz for place of articulation, and 2200 Hz for duration.

This does not mean that under known conditions of distortion one can necessarily predict the individual phonemes which will be heard incorrectly. Rather, this system of feature analysis allows one to predict which phonemes are most likely to be erroneously substituted for the correct phoneme.

The Effects of Hearing Loss on the Perception of Speech Sounds

The attempts thus far to relate a pattern of frequency response in hard-of-hearing listeners to the ability to discriminate a particular consonant have met with little success. It had long been assumed that those individuals with high-frequency hearing losses would have great difficulty in discriminating the high-frequency consonants. Such an assumption seems plausible in the light of the data collected by Fletcher (1929).

Plummer (1943), in his study of 52 hypacusic subjects, failed to find any relationship between acuity losses at high frequencies and difficulty in discriminating the so-called high-frequency consonants. He presented one pair of words at a time, and the listener had to state whether the words were the same or different. No given consonant appeared dependent for its discrimination on any specific frequency. Nor did Schultz (1964) find, for any type of otopathology, any phonemic regularity of consonant errors in response to PB words. Only four errors occurred more than one percent of the time. These were the confusions of /t/ and /k/ with each other and /m/ and /n/ with each other. It should be noted that all four of these are errors involving place of articulation. This high degree of idiosyncracy both within and among pathologic groups weakened the hypothesis that particular phoneme confusions are related to a particular pathology.

Rosen's study (1962) of the responses of 251 sensorineural hypacusics to PB lists indicated that these individuals did not experience as much difficulty in discriminating consonants as might be expected. Fricatives were generally poor, but /ʃ/ and /h/ were among the consonants easiest to discriminate. Nasals were more difficult, and stops, especially the final voiceless ones, were less difficult than had been expected. Rosen also found that the consonants easiest to discriminate were those articulated at or behind the alveolus. In general, the labial and dental consonants were the most difficult. (It is interesting to note that these sounds afford greater visual cues in face-to-face conversation.)

From the data of Schultz and others, it would seem that it might be more rewarding to look for regularity of errors when phonemes are grouped by phonetic features, than

than when they are considered individually. As Schultz points out: ". . . there might be some merit in building an experimental instrument allowing as much interclass response choice as can feasibly be obtained. This, in turn, would allow closer examination of the differences in response class which might differentiate pathologies." (p.11).

A few clinical studies have considered phonetic features. Siegenthaler (1949) studied the relationship between hearing loss and intelligibility of words differing from one another in one of three features. While this study was limited in terms of phonemic inventory and the number of features tested, it did show, for example, that those individuals with high-frequency hearing losses were better at perceiving voicing differences than were those with flat losses. This finding was corroborated by the Miller and Nicely experiment. When only high frequencies are filtered out, the nasality and voicing distinctions are preserved.

When cases of conductive and non-conductive hearing losses were compared, Oyer and Doudna (1959) found more nasality errors for non-conductive losses. However, in general, they did not find great differences between groups in the types of consonant errors made when responding to PB lists.

Although most of these early studies emphasized the role of the consonants, some of the analyses of errors made by hypacusics on PB lists showed that vowel errors were of sufficient frequency to warrant closer investigation. Oyer and Doudna estimated on the basis of frequency of occurrence of vowels and consonants in the lists that the proportion of vowel errors was twice as high as the proportion of consonant errors. Schultz also found that vowel confusions were generally more frequent when errors were converted into percentages.

Oyer and Doudna compared conductive with non-conductive pathologies and found twice the number of vowel confusions among the non-conductives. Schultz compared four types of non-conductive pathologies, finding the same degree of idiosyncrasy within and among groups for the vowels as he had for the consonants. However, differences among groups were evident when the vowel errors were classified by place of articulation. For all pathologies, a prominent finding was the significantly large plurality of front/front vowel substitutions. The other substitution class combinations were less frequent but demonstrated notable differences among groups. The place of articulation confusions for vowels that Schultz reported were in most instances those that would have been predicted from the data of Peterson and Barney (1952) on ease of vowel identification. In addition, the predominance of front/front vowel substitutions is in agreement with Wickelgren's finding of more openness than place errors for vowels.

In light of all the clinical data that have been discussed, it appears that, for both vowels and consonants, specific phoneme substitutions are generally irregular and unpredictable. However, reassessment of errors in terms of phonetic features reveals greater degrees of regularity and predictability of differences among pathologic groups. The clinical findings are generally in agreement with the experimental studies of filtered speech and short-term memory. For example, it has been found in both clinical and experimental studies that those vowels and consonants produced in the front articulatory positions are the most difficult to discriminate.

The apparent relationship between the data derived from clinical and from experimental analyses of the phonetic features of phoneme substitutions warrants further investigation. However, in any further study, a test instrument specifically designed for feature analysis should be employed.

The Development of the Closed-Response Tests

The format of the articulation test appears to be an important factor in phonetic feature analysis. When a closed response set test was used, investigators were better able to control certain relevant factors, and higher significance was generally attributed to the results. Miller (1956), Miller and Nicely (1955), and Siegenthaler (1949) developed special closed-response tests for their studies, and all were able to arrive at certain conclusions regarding the phonetic features involved in response errors. Schultz (1964) also found greater predictability of errors by feature analysis. Yet he concluded that more significant results could have been obtained with a special test and that the PB word lists were inadequate for the purpose of feature analysis.

On open-response tests, such as PB lists, the listener is free to make substitutions for any or all of the phonemes in a word. A confusion of consonants may directly cause a vowel error when the listener substitutes an entirely different word in order to make sense out of his confusion. For example, the stimulus word is "big." The listener confuses the terminal consonant, and hears /k/. Since "bick" is not a familiar word, he may respond with "beak." The examiner will record this as a vowel error as well as a consonant error. Owens *et al*. (1968) suggest that just such a mechanism might account for the high percentage of vowel errors reported by Oyer and Doudna (1959) and by Schultz (1964), both of whom used standard PB lists in their studies. However, unless this mechanism is taken into account, the actual proportion of vowel to consonant errors cannot be known.

Another problem with open-response testing is the lack of control over the available substitutions. The listener may not substitute one particular phoneme for another simply because the word resulting from such a confusion does not exist in the language. For example, if the stimulus word is "cake," the listener is not likely to make an error on the initial phoneme involving the feature of voicing, since "gake" is not an English word. Minimal contrasts of manner of articulation and voicing also result in nonsense words: "hake" and "nake." A minimal contrast of place of articulation is readily available in the familiar word "take." It can be seen from this example that features do not have an equal probability of being confused. Thus, the data obtained from tests such as

PBs may be skewed in the direction of the feature contrast most available among the possible alternatives. Therefore, errors involving all features should be available as choices for each stimulus. This can only be controlled when the entire message set is specified.

Choices are also influenced by the relative distribution of phonemes within the language. The English language, and therefore also PB lists by definition, abound in alveolar sounds. Schultz found that most errors were made on these sounds, and most error responses employed alveolars as substitutions. He suggests that on any new tests non-alveolar sounds be made more available as error choices.

It has also been shown that word familiarity influences intelligibility. If the stimulus is an uncommon word, and the listener is unsure, he is likely to choose a common word similar to the stimulus. Hutton and Weaver (1959) reported that words ranked most familiar on the Thorndike-Lorge lists were significantly more intelligible. Savin (1963) demonstrated that more frequently used words were perceived correctly at much lower signal-noise ratios. Oyer and Doudna estimated that 38 of the 200 PB words account for 50 percent of the total errors. Pollack *et al*. (1959) concluded that PB lists would be adequate for intelligibility testing provided the lists have been practiced by listeners to the point that they approach known message sets. They suggest that the word-frequency effect shown to operate in unknown message sets would not be an important factor were the entire message set known.

One further problem in the use of open-set lists for error analysis is the availability of omissions or additions as erroneous responses. Schultz found that omitted and added phonemes accounted for about half of the total errors. He suggests their elimination as choices on any new tests because they neither yield information about what a patient hears nor differentiate among pathologies.

The major difficulties associated with the use of open-response test materials, such as PB lists, in the analysis of phoneme substitutions may be summarized as follows: lack of control over the phoneme being tested and the kinds of alternatives available, and the effects of word and phoneme frequency. One technique that minimizes all of these problems is the closed-response set.

Many of the recently designed closed-response tests are modifications of the Fairbanks Rhyme Test (1958). The complete Rhyme Test is made up of 250 common monosyllables which form 50 sets of five rhyming words each. The subject's response sheet shows 50 words stems, that is, rhyming portions that remain when the initial consonants are omitted. Each 50-word list contains one word from each set of five rhyming words. The subject's task is to fill in the initial consonant, in each case, a single letter. Thus, the same response sheet serves for each of the five lists. This tests meets the need of testing only one phoneme at a time. However, since the exact message set is not specified for the listener, the range of alternatives consists of all words with a given stem, regardless of phonemes, features, or word frequency.

The Fairbanks test was modified by House *et al.*(1965). The *Modified Rhyme Test* (MRT) differs from the Fairbanks version in that the available responses are defined for the listener. For each test item, the listener selects the word he hears from the six alternatives on the answer sheet. Each of the alternatives serves as the test item on one of the six resultant lists. Many of the response sets are made up of words which present mostly multi-feature contrasts. House admits that his material requires retabulation in terms of ponetic features, since the response choices were not arranged to provide distinctive contrast of features.

In the interest of testing a greater number of minimal-feature contrasts, Griffiths (1967) revised the *MRT*. Those sets of words which consisted of multi-feature contrasts, as well as those sets which repeated minimal contrasts found in other sets, were replaced by sets that tested minimal contrasts not present in the House lists. Griffiths believed that a test which revealed all the basic one-feature confusions could validly predict multi-feature confusions as well. Griffiths used three features, manner, voicing, and place to classify the consonants. He indicated that there were several one-feature contrasts in this test which would be considered multi-feature contrasts when described by the system of Chomsky and Halle.

Griffiths' test is made up of 50 sets of five alternatives. The multiple-choice format used by House was retained. Each of the alternatives serves as the test item on one of the five lists. Griffiths analyzed the confusions which occurred when the lists were presented at various signal-noise ratios. Errors were classified as either errors in one feature only, or as multi-feature errors. When an error involved two or more features, examination of the particular words in the set showed that the chosen phoneme was the only one available which provided a certain feature contrast. Griffiths assumed that such choices resulted from the shift of some perceptual cue in one independent feature, and therefore, that the choice of a word differing in two features could be interpreted as an error in one of them. The remaining effective multi-dimensional errors were 0 percent at +30 dB S/N, 2 percent at +8 S/N, and 3 percent at +4 S/N. It was concluded that the minimal feature contrasts upon which the test was based are independent features. It was also found that, with this type of closed-set response format, there was no indication of any effect due to unfamiliar words or pronunciations.

Pickett _et_ _al_. (1972) constructed a special version of the Fairbanks-House Modified Rhyme Test. They tested 20 initial consonants, 20 final consonants, and 10 vowels and diphthongs. Each test item consisted of six monosyllabic response words, and the stimulus word was chosen at random from the six response words. The test was administered to 99 subjects with moderate or severe sensorineural hearing losses. Responses were analyzed by examining individual vowel confusions and by calculating the proportions of consonant confusions involving voicing, manner, and place of articulation. Results showed that vowels were far more intelligible than consonants. However, when open-set PB responses were analyzed, the opposite order of intelligibility was found: Oyer and Doudna (1959) and Schultz (1964) suggested that vowels were _less_ intelligible than consonants.

Pickett _et al_. also found that consonants were more intelligible in word-initial than in word-final position. Analysis of feature confusions showed a majority of place confusions. In addition, they reported fewer voicing confusions for final than initial consonants. Pickett _et al_. suggest that voicing of final consonants may be perceived by discriminating vowel durations. They concluded that sensorineural listeners have better residual reception for low-frequency speech features, such as voicing, nasality, and first formants of vowels, than for higher-frequency features, such as consonant noise bursts and friction, and higher vowel formants.

Studebaker (1968) also modified the procedure of House _et al_. A four-alternative "wide-range" consonant test was developed. Each set consisted of two pairs of words. The words in each pair differed from each other in place of articulation, while the pairs differed from each other in terms of manner, voicing, or nasality. For example, one set contains two voiceless stops and two voiceless fricatives. The lists were presented to hard-of-hearing listeners. It was found that even those subjects with poor discrimination could easily categorize words as voiced, voiceless, or as nasal. It was decided to use this test with children having severe discrimination problems. For adults, a new test concentrating on place of articulation was devised.

The "narrow-range" test, which Studebaker then developed, consisted of sets requiring one of four choices, namely, one of three consonants differing in place of articulation only, or complete omission of the consonant. Since only one feature was being tested, the results were presented in terms of phoneme-for-phoneme substitutions. As would be expected from previous studies and observations, the most common error was the /f/-/θ/confusion; the second most common error was the /v/- /ð/. Omissions were often chosen for /v/ or /ð/, but rarely for /f/ or/θ/.

House, Griffiths, Pickett _et al_., and Studebaker all tested final as well as initial consonants. Studebaker found that consonants in the initial position were generally easier to identify than those in the final position. This was true for hypacusics, as well as for normal hearers given filtered speech. This is consistent with the findings of House _et al_. and Pickett _et al_., but not, however, with Griffiths' finding of no significant difference between the number of initial and final consonant errors.

Studebaker also presented a vowel test consisting of eight vowels in an /s-t/ environment. For each vowel presented, the other choices consisted of the remaining seven. For both normals listening under various S/N ratios and the hard-of-hearing subjects, certain patterns were significant. For both groups, the most common errors were /ɪ/ for /ʊ/, and /æ/ for /ɛ/. It would seem possible that the first of these confusions is explainable on the basis of second formant confusion, since these are quite different, while the first formants are quite similar. However, the reciprocal error, /ʊ/ for /ɪ/ rarely occurred. In addition, when considering the /æ/ for /ɛ/ confusion, comparison of second formants is no more enlightening than comparison of first formants. In general, it seems that each vowel presented a unique problem to the subject. Studebaker's data suggest that his tests are better able than are pure-tone or PB tests to rank the hard-of-hearing on their discrimination abilities.

For error analysis, closed-response set articulation tests have been shown to have several advantages. Written multiple-choice tests allow a degree of ease and accuracy of scoring which cannot be obtained when the examiner must record the verbal responses of the listeners. Controls can be exerted over the phoneme under test and the available alternatives. Such controls are necessary if phoneme substitutions are to be analyzed or if the proportions of errors for initial consonants, final consonants, or vowels are to be calculated. The effects of learning, word familiarity, and other response bias can be minimized without the need for extensive prior training.

In the last decade, several closed-response set tests have been developed. In general, they have been shown to be effective diagnostic tools in the systematic analyses of discrimination of speech sounds.

Summary

It has been demonstrated that one cannot predict an individual's speech discrimination ability from other audiometric measures. Direct tests of speech discrimination were developed on the principle of phonetic balance, which, till recently, has been the most widely researched aspect of articulation testing.

The contributions of various portions of the speech spectrum to speech perception have been studied extensively. The discrimination of speech sounds is dependent on several kinds of acoustic cues. These cues have been shown to be often, but not always, frequency-dependent. Timing factors have also been shown to play an important role in speech perception. An efficient way of describing a phoneme is to consider it, not as a single unit, but rather as a set of phonemic features.

It has been shown clinically and experimentally that these phonemic features operate somewhat independently. One or more features can be perceived or recalled, when others cannot. The effect of various kinds of distortion on the integrity of these features can be a diagnostically valuable test. However, the nature of the test instrument is critical in feature analysis. What is needed is a closed-response set test specifically designed to allow as many direct minimal feature contrasts as possible, with a minimum of of subjective interpretation of data.

CHAPTER III

DEVELOPMENT OF THE PHONEME DIFFERENTIATION TEST

The closed-response Phoneme Differentiation (PD) Test was developed in accordance with the criteria outlined in Chapter I of this paper. The 200-item list allows as many minimal feature contrasts per item as could be developed within the limits of the English language. The contrasts are based on a model of phonemic features adapted from existing models. Both vowels and consonants appear as test phonemes.

The word list was recorded and presented unfiltered to hard-of-hearing listeners, and under various conditions of low-pass filtering to normal-hearing listeners. The latter group would make very few errors unless the stimuli were distorted in some way. Removal of various amounts of high frequencies result in various degrees of intelligibility, as shown in the Miller and Nicely study. In other words, low-pass filtering is one way to drive scores low enough so that a sufficient number of errors are provided for analysis. This particular form of signal distortion was chosen because it relates to the high-frequency hearing losses of many hypacusics.

The errors for each phonemic feature, assigned in advance to each response, were tabulated for each listener and each listening condition.

The Word List

One of the major differences between the present set of materials and previously developed tests is that, for each test word, a special set of alternatives was devised. Thus each set of words, or test item, is made up of a test phoneme and four alternatives, each of which presents a minimal feature contrast to the test phoneme. Each set of words can be used only once in the test. The entire message set, five choices for each test item, is available to the listener on the response sheets provided. The task of the listener is to circle the word that he hears among the choices available to him.

For each set of words comprising one test item, the variation occurs in one phoneme at a time. For some sets, the variation is the consonant in the word-initial position, while the following vowel and consonant remain constant. For other sets, the variation is the consonant in the word-terminal position. For others, only the syllable nucleus (vowel or diphthong) varies. All of the test words and their alternatives are CVC. Therefore, when a consonant is the test phoneme, it is always adjacent to a vowel, never another consonant. When a vowel is the test phoneme, it is always preceded and followed by a consonant.

Thus, as with the tests described by House *et al.* and Griffiths, only one phoneme is tested at a time, the choices are controlled, and omissions and additions can be eliminated as responses.

One of the objectives of the present study is to avoid individual interpretation of features involved in error responses, as is needed for the House *et al*. and Griffiths tests. I, therefore, decided to follow a model of phonemic features in the development of the word sets. Each choice for each word represents an error in a particular feature, and, wherever possible, the difference is in that particular feature only.

Consonants

The phonemic model chosen for the consonants was adapted from that of Wickelgren (1965b), which was shown to be the most accurate in prediction of errors in short-term memory. Four dimensions or features are used to describe consonant phonemes. These features have been retained in the present model, but the assignment of values to the phonemes differs somewhat. These features, with their corresponding values, are presented in Table 7 and are described below.

Voicing.- This feature has two values, voiced or voiceless. Although voicing is acoustically a continuum from voicing lead, through simultaneous onset, to voicing lag, there are, in English, only two values perceptually (Lisker and Abramson, 1964, 1967). To attempt further description within a phonemic model would be complicated by the variations for different speakers and different phonetic

TABLE 7

THE FEATURE SYSTEM FOR CONSONANTS USED IN THE PHONEME DIFFERENTIATION (PD) TEST

Consonant	Voicing	Nasality	Openness	Place
* /p/	0	0	0	0
* /b/	1	0	0	0
* /m/	1	1	0	0
* /t/	0	0	0	2
* /d/	1	0	0	2
* /n/	1	1	0	2
/tʃ/	0	0	0	3
/dʒ/	1	0	0	3
* /k/	0	0	0	4
* /g/	1	0	0	4
** /ŋ/	1	1	0	4
* /f/	0	0	1	0
* /v/	1	0	1	0
/θ/	0	0	1	1
* /ð/	1	0	1	1
* /s/	0	0	1	2
* /z/	1	0	1	2
/ʃ/	0	0	1	3
** /ʒ/	1	0	1	3
/h/	0	0	1	4
/w/	1	0	2	0
/l/	1	0	2	2
/j/	1	0	2	3
/r/	1	0	2	4

* These consonants are tested in both word-initial and word-terminal position.

** Although these consonants are not presented as test items, their classifications are given here because they are used as alternatives to other test phonemes.

environments. This feature is the same for all feature models described in this paper. Although in some models it may be called "tense-lax" instead of "voiced-voiceless," the difference is only in the label or in the theory behind it. The phonemes are sorted into identical groups.

Nasality.- This feature also has two phonemically relevant values, present or absent, and like the voicing feature, is the same for all of the phoneme models.

Openness of the vocal tract.- This dimension handles what is usually termed "manner of articulation." It incorporates the affrication and duration features of the Miller and Nicely model. There are three values indicating the degree of constriction in the vocal tract during the production of the consonant:

(0) includes the stops, affricates, and nasals, which are all produced with a complete obstruction in the oral cavity. The nasals are classed as stops in terms of their oral articulations. The additional open nasal cavity is handled by the feature "nasality."

(1) includes the fricatives, which are produced with a constriction through which air is forced, generating sound by friction.

(2) includes the semivowels /w,r,l,j/, which are produced with less constriction than the other consonants.

Place of articulation.- This dimension differs slightly from Wickelgren's model in terms of the assignments of values to the phonemes. Five values from front to back, designated by the numerals 0-4, are necessary to separate those phonemes with the same values for the other three features. Some of these values encompass two adjacent places of articulation, such as bilabial and labiodental. However, further division was not considered necessary, since the phonemes still have unique descriptions, and the places would have to be merged in the applications of the model to word contrasts.

(0) includes bilabials and labiodentals;
(1) includes linguadentals;
(2) includes alveolars;
(3) includes palatals;
(4) includes velars and glottals.

This place feature with multiple values was chosen in light of previous experimental evidence (Wickelgren, 1965b; Klatt, 1968). However, the feature is tested only once for each test item. Since the place feature is multivalued, a choice had to be made in each case as to which place value would be presented in contrast to the value for the test phoneme. For different test items involving the same consonant, at least two different contrasting place values were presented. For example, of the four test items for initial /p/, /t/ was given in two cases, and /k/ was given in the other two cases. More than one place contrast per set would allow a higher probability of place errors on the test.

The purpose of the phonemic model given here is to provide a unique, though not complete, description of each phoneme. Certain values for some features are grouped, resulting in losses of some distinctions. For example, both labiodental and bilabial places of articulation are assigned the same value. This distinction is not essential in English since there are neither labiodental plosives nor bilabial fricatives in the language.

The basic design of the consonant test requires the opportunity for the choice of an error in one feature at a time. The listener can choose the correct word, or a word differing from it in one of four features: voicing, nasality, openness or place. For example, when the test phoneme is initial /p/, and the test word is "pill," "bill" presents a voicing contrast, "mill" presents a nasality contrast, "till" presents a place contrast, and "fill" presents a contrast in openness of the vocal tract.

One of the above contrasts, "mill," is not truly minimal, that is, it presents contrasts in more than one feature. However, since there are no voiceless nasal phonemes, this contrast is inherently multi-dimensional. In addition, it can be assumed from the outset that the choice of "mill" for "pill" would represent an error in nasality, since an error in voicing alone, "bill," is also available. This same procedure must also apply to the semivowels, since they are also without voiceless phonemic cognates. Thus voicing contrasts for the purposes of the test are presented in voiceless fricatives. Other exceptions to the minimal feature contrast criterion are the openness and voicing contrasts for the nasals. In these cases the nasals are treated like voiced stops produced in the same place of articulation.

The phonemes for which appropriate test words and alternates could be developed include twenty-two consonants in the word-initial position and thirteen in the word-terminal position. The /ʒ/ and /ŋ/ are the only consonants not tested, because they do not occur initially in English, and appropriate sets of foils could not be found for them in the word-terminal position.

The consonants tested in both word-initial and word-terminal positions include /p,t,k,b,d,g,m,n,f,s,z,ð/. /w,r,h,j/ were not tested in the word-terminal position because they do not exist as separate consonant phonemes in that position. The phoneme /l/ was not tested in the terminal position because the other semivowels were not available for place contrasts. The phonemes /tʃ, dʒ,θ,ʃ/ were not tested in the terminal position because enough appropriate foils could not be found for the test words containing these phonemes.

Vowels

As is evident from the feature systems described in Chapter II of this paper, the set of features necessary for a description of the vowels differs from the set of features used to described the consonants. Two of the features for consonants are inappropriate for vowels, since these phonemes do not have voiceless or nasal cognates on the phonemic level in English. However, two of the consonant features have the same relationship to the vowels.

The openness feature, which describes the degree of constriction of the vocal tract during consonant production, is the same in acoustic and articulatory terms as the height of the tongue during vowel production. When all phonemes are described by this feature, the vowels are assigned higher values of openness than the consonants. Thus, the order of degree of openness for all classes of phonemes would be, in increasing order: stops, fricatives, semivowels, close vowels, medium vowels, and open vowels. Acoustically, this order of phonemic class also represents a progressive increase in frequency of the first formant.

The place of articulation feature is also similar for consonants and vowels. Just as consonants are described by the point along the vocal tract of maximum constriction, the vowels are described by the point along the vocal tract of maximum tongue height. In addition, the progression from back to central to front vowels exhibits a corresponding rise in the frequency of the second formant.

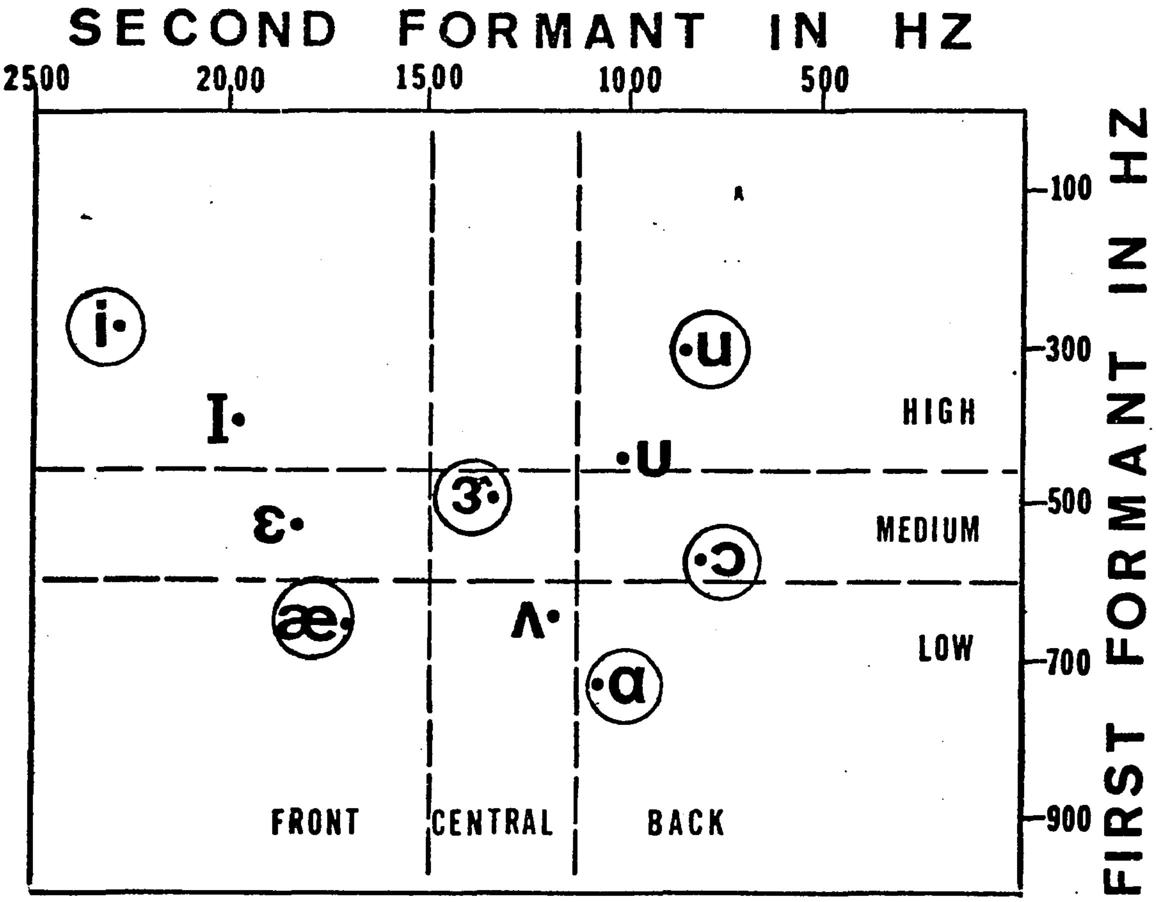
As seen from the measurements of formant frequencies for many speakers (Peterson and Barney, 1952), a plot of the first two formants is not sufficient for a unique vowel description. The third formant frequency would serve to separate overlapping vowels.

Description by distinctive features requires that absolute frequency values be replaced by classification into a limited number of categories, even if these categories can be defined by frequency. Both phonemic description and formant frequency can be considered simultaneously in the manner shown in Figure 1. The ten vowels used in the present study may be classified as close, medium or open by grouping them according to a certain range of first formant values (shown on the vertical axis). In the same manner, they may be classified as front, central, or back according to certain ranges of second formant values (shown on the horizontal axis).

Of the nine cells which result, one is empty, six have one vowel each, and two have two vowels each. In order to provide a completely unique description, that is, only one vowel per cell, one would need further divisions along one of the axes, resulting in a greater number of values for one of the formants. However, it was felt that division of these two features into a greater number of values might be somewhat artificial, in view of the fact that there is overlap between vowels in the ranges of formant frequencies. In addition, the decision to further divide a cell with two vowels, such as the one containing /i/ and /ɪ/, would be arbitrary, since these vowels show no greater difference in one formant than in another when plotted on the mel scale as shown by Peterson and Barney (1952).

Consideration of the third formant would separate

Figure 1. Frequencies of the first and second formants and phonetic descriptions of ten vowels. Long vowels are circled. (Formant values are those found by Peterson and Barney, 1952)


SECOND FORMANT IN HZ
2500
2000
1500
1000
500
FIRST FORMANT IN HZ
100
300
500
700
900
HIGH
MEDIUM
LOW
FRONT
CENTRAL
BACK
i
ɪ
ɛ
æ
ɝ
ʌ
u
ʊ
ɔ
ɑ

/i/ from /ɪ/, since the difference in third formant frequencies is approximately 500 Hz. However, it is less certain that the third formant frequency difference of about 50 Hz or less would serve to separate /ʊ/ from /u/. In view of the intended and probable uses of the present test materials, under conditions of noise, low-pass filtering, and with hard-of-hearing listeners, the use of the third formant as a distinctive feature appears irrelevant, since it is not likely to be transmitted.

Since it has been shown that duration operates as a cue in vowel perception (Miller, 1956), it was decided to use it as a third feature for vowels. Inspection of Figure 1 shows that for those cells containing two vowels, one of each pair is long--/i,u/, and one is short--/ɪ,ʊ/.

Thus three features which can provide a unique description for each of the ten vowels considered here are: openness, place of articulation, and duration. These features and their corresponding values are shown in Table 8.

The application of the vowel model to the choices of alternate words is not a clear-cut procedure as is the application of the consonant model. The selection of contrasts is illustrated in Table 9. The contrast selected for each vowel for each feature is that vowel which differs in the feature in question but does not differ in the other two features. For example, a minimal openness contrast for the vowel /i/ (openness value =0) is provided by the vowel /æ/ (openness value =2). Both /i/ and /æ/ have a place

TABLE 8

THE FEATURE SYSTEM FOR VOWELS USED IN THE PHONEME DIFFERENTIATION TEST

Vowel	Openness	Place	Duration
/i/	0	0	1
/ɪ/	0	0	0
/ɛ/	1	0	0
/æ/	2	0	1
/ɝ/	1	1	1
/ʌ/	2	1	0
/ɑ/	2	2	1
/ɔ/	1	2	1
/ʊ/	0	2	0
/u/	0	2	1
	0=close 1-medium 2=open	0=front 1=central 2=back	0=short 1=long

TABLE 9

AVAILABLE ONE-FEATURE CONTRASTS FOR VOWELS

Vowel	Feature Description[a]			Openness Contrast	Place Contrast	Duration Contrast
	Open.	Place	Duration			
/i/	0	0	1	(2)01/æ/	0(2)1/u/	00(0) /ɪ/
/ɪ/	0	0	0	(1)00/ɛ/	0(2)0/ʊ/	00(1) /i/
/ɛ/	1	0	0	(0)00/ɪ/	1(-)0	10(-)
/æ/	2	0	1	(0)01/i/	2(2)1/ɑ/	20(-)
/ɝ/	1	1	1	(-)11	1(2)1/ɔ/	11(-)
/ʌ/	2	1	0	(-)10	2(-)0	21(-)
/ɑ/	2	2	1	(1)21/ɔ/ (0)21/u/	2(0)1/æ/	22(-)
/ɔ/	1	2	1	(2)21/ɑ/ (0)21/u/	1(1)1/ɝ/	12(-)
/ʊ/	1	2	0	(-)20	0(0)0/ɪ/	02(1) /u/
/u/	0	2	1	(1)21/ɔ/ (2)21/ɑ/	0(0)1/i/	02(0) /ʊ/
Total Number of Contrasts Available				7/10	8/10	4/10

[a] From Table 8.

value of <u>0</u>, and both have a duration value of <u>1</u>.

Inspection of Table 9 reveals that, of the thirty possible one-feature contrasts (three for each of ten vowels), nineteen exist in the language according to the model as described. Therefore, in applying the model to the word list, other vowel foils had to be chosen which differed in the described feature, but were as similar as possible in the other two features.

The fourth alternative selected for each vowel was one which differed from the test vowel in as many features as possible. In all cases this alternative involved contrasts in at least two features simultaneously. In some cases, the contrasts were in openness and place; in some, they were in openness and duration; in others, they were in place and duration; and in most cases, the fourth alternative contrasted along all three dimensions.

This multiple-feature contrast deviates from the minimal-feature criterion. However, it was selected for several reasons. First, in order to keep the probability of error equal for both vowels and consonants, all test sets had to have the same number of alternatives available to the listener. Secondly, a fourth feature could not be found which would not overlap considerably with the three features already chosen. Thirdly, the relatively easy discrimination between vowels differing in more than one feature would provide a reasonable parallel to the relatively easy discrimination between consonants differing in the voicing feature alone.

The analysis of vowel errors will necessarily be on an individual basis. The evidence for actual minimal-feature contrasts of vowels must await analysis of the data.

Diphthongs

In the same format as for vowels, five diphthongs were tested. Each diphthong was given the opportunity to be confused with as many of the other four diphthongs as possible. The remaining contrasts were mostly long vowels, /i,æ,ɝ,ɑ,ɔ,u/, which had formants similar to the onset or offset of the diphthong being tested.

The complete test of initial consonants, terminal consonants, vowels and diphthongs is given in Tables 10, 11, 12, and 13, respectively.

For both consonants and vowels, an attempt was made to vary the phonetic environments. Where possible, the same consonant was placed before or after a front vowel in one token, a central vowel in another, and a back vowel in another. The vowels and diphthongs, on the other hand, were placed between consonants which were in some tokens front, in others middle, and in others back.

Another factor considered was word familiarity. Test words and their alternatives were checked for frequency of usage on the Thorndike-Lorge Lists of frequently-used words (1944). More familiar words were used where possible.

TABLE 10

INITIAL CONSONANT *PD* TEST ITEMS

Phoneme	Test Word	Voicing Contrast	Nasality Contrast	Openness Contrast	Place Contrast
/p/	PAT	BAT	MAT	FAT	CAT
	PAN	BAN	MAN	FAN	TAN
	PILL	BILL	MILL	FILL	TILL
	PAUL	BALL	MALL	FALL	CALL
/t/	TILL	DILL	NIL	SILL	KILL
	TOT	DOT	NOT	SOT	POT
	TIP	DIP	NIP	SIP	KIP
	TORE	DOOR	NOR	SORE	PORE
/k/	COT	GOT	NOT	HOT	POT
	KILL	GILL	NIL	HILL	TILL
	CAT	GAT	NAT	HAT	PAT
	COAL	GOAL	KNOLL	HOLE	TOLL
/tʃ/	CHILL	JILL	NIL	SHILL	KILL
	CHEST	JEST	NEST	FEST	TEST
	CHAR	JAR	MAR	FAR	TAR
	CHEWED	JUDE	MOOD	SUED	COOED
/b/	BET	PET	MET	VET	GET
	BAIL	PAIL	MAIL	VEIL	DALE
	BOWL	POLE	MOLE	VOLE	GOAL
	BORE	PORE	MORE	FORE	DOOR
/d/	DIE	TIE	NIGH	THY	BUY
	DAB	TAB	NAB	LAB	GAB
	DEED	TEED	NEED	LEAD	BEAD
	DO	TO	NEW	ZOO	GOO
/g/	GILL	KILL	NIL	WILL	BILL
	GALE	KALE	NAIL	WAIL	DALE
	GOT	COT	NOT	WATT	DOT
	GOAL	COAL	KNOLL	ROLL	BOWL
/dʒ/	GYP	CHIP	NIP	RIP	DIP
	JET	CHET	NET	YET	BET
	JAR	CHAR	MAR	CZAR	BAR
	JUDE	CHEWED	NUDE	RUDE	DUDE

TABLE 10 CONTINUED

Phoneme	Test Word	Voicing Contrast	Nasality Contrast	Openness Contrast	Place Contrast
/m/	MILL	PILL	BILL	WILL	NIL
	MAIL	PAIL	BAIL	VEIL	NAIL
	MUTT	PUTT	BUT	RUT	NUT
	MORE	PORE	BORE	WORE	NOR
/n/	NEAR	TIER	DEAR	LEER	MERE
	NAIL	TAIL	DALE	YALE	MAIL
	KNOCK	TOCK	DOCK	LOCK	MOCK
	NEW	TOO	DO	ZOO	MOO
/f/	FEEL	VEAL	MEAL	PEAL	SEAL
	FEAR	VEER	MERE	PIER	SHEER
	FINE	VINE	MINE	PINE	SHINE
	FAULT	VAULT	MALT	PALT	SALT
/θ/	THINK	ZINC	MINK	PINK	SINK
	THICK	VIC	NICK	PICK	SICK
	THIGH	THY	MY	TIE	SHY
	THOUGHT	BOUGHT	NOUGHT	TAUGHT	SOUGHT
/s/	SINK	ZINC	MINK	CHINK	FINK
	SAD	LAD	MAD	TAD	SHAD
	SOUGHT	BOUGHT	NOUGHT	TAUGHT	FOUGHT
	SOWN	ZONE	KNOWN	TONE	SHOWN
/ʃ/	SHIP	ZIP	NIP	KIP	SIP
	SHIN	GIN	MIN	KIN	FIN
	SHOWN	ZONE	KNOWN	CONE	PHONE
	SHOES	ZOOS	NEWS	TWOS	SUES
/h/	HEAT	WHEAT	NEAT	KEET	SHEET
	HAT	THAT	NAT	CAT	FAT
	HUM	RUM	NUMB	COME	THUMB
	HOLE	GOAL	KNOLL	COAL	SOLE
/v/	VEST	FEST	NEST	BEST	ZEST
	VEIL	FAIL	MAIL	BAIL	THEY'LL
	VINE	FINE	MINE	BINE	THINE
	VAT	FAT	MAT	BAT	THAT
/ð/	THEN	HEN	MEN	DEN	ZEN
	THINE	FINE	MINE	DINE	VINE
	THY	THIGH	MY	BY	VIE
	THAT	FAT	MAT	BAT	VAT

TABLE 10 CONTINUED

Phoneme	Test Word	Voicing Contrast	Nasality Contrast	Openness Contrast	Place Contrast
/z/	ZEAL	SEAL	KNEEL	DEAL	VEAL
	ZED	SAID	NED	DEAD	JED
	ZONE	SOWN	KNOWN	LOAN	JOAN
	ZOO	SUE	NEW	DO	VOO
/w/	WENT	SENT	MEANT	BENT	LENT
	WAIL	FAIL	MAIL	BAIL	YALE
	WIND	FIND	MIND	BIND	RIND
	WORE	FORE	MORE	BORE	ROAR
/r/	RIP	SIP	NIP	GYP	LIP
	RATE	SATE	MATE	DATE	WAIT
	ROT	SHOT	NOT	JOT	YACHT
	RUES	SHOES	NEWS	ZOOS	WOOS
/l/	LEAP	SEEP	NEAP	DEEP	WEEP
	LIP	SIP	NIP	DIP	RIP
	LOT	SOT	NOT	DOT	ROT
	LOSE	SUES	NEWS	ZOOS	USE
/j/	YEAR	SHEER	NEAR	JEER	REAR
	YACHT	SHOT	NOT	GOT	LOT
	YALE	SHALE	NAIL	JAIL	WAIL
	YOU	SHOE	NEW	GOO	WOO

TABLE 11

TERMINAL CONSONANT PD TEST ITEMS

Phoneme	Test Word	Voicing Contrast	Nasality Contrast	Openness Contrast	Place Contrast
/p/	PUP	PUB	PUN	PUFF	PUCK
	ROPE	ROBE	ROAM	ROACH	WROTE
	CUP	CUB	COME	CUFF	CUT
	RIP	RIB	RIM	RIFF	RICK
/t/	KIT	KID	KIN	KISS	KICK
	LIGHT	LIED	LINE	LICE	LIKE
	SIT	SID	SIN	SIS	SIP
	COAT	CODE	CONE	COAL	COPE
/k/	LEAK	LEAGUE	LEAN	LEASH	LEAP
	WICK	WIG	WING	WISH	WIT
	BUCK	BUG	BUNG	BUS	BUT
	COKE	CODE	CONE	COACH	COPE
/b/	ROBE	ROPE	ROAM	ROVE	RODE
	COB	COP	CON	COL	COD
	RIB	RIP	RIM	RILL	RIG
	DUB	DUCK	DUMB	DOVE	DUG
/d/	WED	WET	WHEN	WELL	WEB
	BUD	BUT	BUN	BUZZ	BUG
	RID	WRIT	RIM	RILL	RIG
	ROAD	WROTE	ROAN	ROSE	ROBE
/g/	BUG	BUCK	BUNG	BUZZ	BUD
	DUG	DUCK	DONE	DOES	DUB
	LEAGUE	LEAK	LEAN	LEES	LEAD
	LAG	LACK	LAMB	LALL	LAB
/m/	ROAM	ROPE	ROBE	ROVE	ROAN
	RIM	RIP	RIB	RILL	RING
	CAME	CAPE	CAGE	CAVE	CANE
	SUM	SUP	SUB	SULL	SUNG
/n/	BUN	BUT	BUD	BUZZ	BUNG
	KIN	KIT	KID	KILL	KING
	MAIN	MATE	MADE	MAZE	MAIM
	ROAN	WROTE	ROAD	ROSE	ROAM

TABLE 11 CONTINUED

Phoneme	Test Word	Voicing Contrast	Nasality Contrast	Openness Contrast	Place Contrast
/f/	LEAF	LEAVE	LEAN	LEAP	LEASH
	REEF	REAVE	REAM	REAP	WREATH
	LIFE	LIVE	LIME	LIGHT	LICE
	ROOF	ROOVE	ROOM	ROOT	RUTH
/s/	LICE	LIES	LINE	LIGHT	LIFE
	BUS	BUZZ	BUN	BUT	BUFF
	LOOSE	LOSE	LOOM	LOOT	LOOF
	LEASE	LEES	LEAN	LEAK	LEASH
/ð/	TEETHE	TEETH	TEAM	TEED	TEASE
	SOOTHE	SOOTH	SOON	SUED	SUES
	LITHE	LIFE	LIME	LIED	LIVE
	LOATHE	LOAF	LOAN	LOBE	LOAVE
/v/	LIVE	LIFE	LIME	LIED	LITHE
	WAVE	WAIF	WANE	WADE	WAYS
	DOVE	DUFF	DUMB	DUB	DOES
	LOAVE	LOAF	LOAM	LOBE	LOATHE
/z/	SEIZE	CEASE	SEEN	SEED	SEETHE
	LIES	LICE	LINE	LIED	LIVE
	DOZE	DOSE	DOME	DOLE	DOGE
	BUZZ	BUS	BUN	BUD	BOVE

TABLE 12

VOWEL PD TEST ITEMS

Phoneme	Test Word	Place Contrast	Openness Contrast	Duration Contrast	Multiple Contrast	Number of Multiple Feature Differences
/i/	KEEP	COOP	CAP	KIP	COP	2
	TEAM	TOMB	TAM	TIM	TOM	2
	LEAK	LUKE	LACK	LICK	LOOK	2
	PEAT	PUT	PAT	PIT	POT	2
/ɪ/	BID	BUD	BED	BEAD	BAD	2
	SIP	SUP	SAP	SEEP	SOUP	2
	KIN	KERN	KEN	KEEN	CON	3
	BIG	BUG	BEG	BERG	BOG	3
/ɛ/	PECK	PUCK	PICK	PEAK	PERK	2
	NET	NUT	KNIT	NEAT	NOT	3
	CHECK	CHUCK	CHICK	CHEEK	CHOCK	3
	PEP	PUP	PIP	PEEP	POP	3
/æ/	BAG	BOG	BIG	BUG	BERG	2
	PAT	POT	PEAT	PET	PUT	3
	SAP	SOP	SEEP	SUP	SOUP	2
	CAT	COT	KEET	CUT	COOT	2
/ɝ/	TERM	TOMB	TOM	TIM	TEAM	2
	BIRD	BID	BAD	BUD	BEAD	2
	LURK	LUKE	LOCK	LUCK	LEAK	2
	HURT	HOOT	HOT	HUT	HEAT	2
/ɑ/	TOM	TAM	TOMB	TUM	TIM	3
	COP	CAP	COOP	CUP	KIP	3
	LOCK	LACK	LUKE	LUCK	LICK	3
	POT	PAT	PERT	PUTT	PIT	3
/ɔ/	DAWN	DAN	DUNE	DONE	DIN	3
	CAWED	CAD	COOED	COULD	KID	3
	BOUGHT	BAT	BOOT	BUT	BIT	3
	TALK	TACK	TOCK	TUCK	TICK	3

TABLE 12 CONTINUED

Phoneme	Test Word	Place Contrast	Openness Contrast	Duration Contrast	Multiple Contrast	Number of Multiple Feature Differences
/ʌ/	DUCK	DECK	DICK	DOCK	DUKE	3
	BUT	BET	BOOT	BERT	BEET	3
	PUN	PEN	PIN	PAWN	PAN	2
	CUD	KED	COULD	CURD	KEYED	3
/ʊ/	COULD	KID	COD	COOED	CAD	3
	SOOT	SIT	SOT	SUIT	SAT	3
	PUT	PIT	POT	PEAT	PAT	3
	HOOK	HICK	HOCK	HAWK	HACK	3
/u/	BOOT	BEET	BOUGHT	BUT	BET	3
	DUNE	DEAN	DAWN	DONE	DEN	3
	COOP	KEEP	COP	CUP	KIP	2
	LUKE	LEAK	LOCK	LOOK	LICK	2

TABLE 13

DIPHTHONG *PD* TEST ITEMS

Phoneme	Test Word				
/aɪ/	RISE	ROY'S	ROUSE	ROSE	RAISE
	NINE	NOON	NOUN	KNOWN	NAN
	PIES	POISE	PAYS	POSE	PAUSE
	TIDE	TOYED	TOAD	TEED	TAD
/eɪ/	PAYS	POISE	POSE	PIES	PEAS
	BAIT	BITE	BOAT	BOUT	BEET
	TALE	TOIL	TOLL	TOWEL	TILE
	GATE	GOT	GOAT	GOUT	GERT
/ɔɪ/	LLOYD	LIED	LOAD	LOUD	LEAD
	POISE	PIES	POSE	PAYS	PEAS
	BOYS	BUYS	BEAUS	BOUGHS	BOOZE
	TOYED	TIDE	TOAD	TOD	TEED
/aʊ/	BOUT	BOAT	BITE	BAIT	BOOT
	LOUD	LOAD	LIED	LLOYD	LEWD
	FOUND	PHONED	FIND	FEIGNED	FIEND
	SOWS	SEWS	SIZE	SAWS	SEES
/oʊ/	BEAUS	BOUGHS	BOYS	BUYS	BAYS
	BOAT	BOUT	BITE	BAIT	BOOT
	ROSE	ROUSE	ROY'S	RISE	RAISE
	TONE	TOWN	TINE	TURN	TEEN

However, the criterion of minimal feature contrasts was not sacrificed for the sake of word familiarity, and it was felt that it would not be a critical factor in any event, since the entire message set is presented to the listener. The infrequently used words that do remain are simple CVC monosyllables, the pronunciation of which can be easily determined from their spellings and rhyme-set environments. Such words include: "riff, rill, reave, roove," and "sull." There are also a few nonsense syllables on the list. These are "palt, bove, voo," and "loave," and they appear to be equally easy to pronounce for the same reasons.

Each of the fifty phonemes, including twenty-two initial consonants, thirteen terminal consonants, ten vowels, and five diphthongs, are tested four times each. The complete list of two hundred words might thus be divided into four lists, each of which tests the same fifty phonemes. However, before such division can be undertaken, the entire list must be analyzed for comparability of different tokens and the effects of phonetic environment.

Randomization and test forms

In the first test order, randomized tokens of initial consonant test items are presented first, followed by randomized vowel and diphthong test items, followed by randomized terminal consonant test items. It was felt that randomization of the list as a whole would complicate the task for the listener. For each item presented, one would have to switch from listening for initial consonant differences to vowel differences, and so forth. A second randomization was made beginning with vowel and diphthong items, followed by terminal consonants, followed by initial consonants. These two randomizations are referred to as Test Orders 1A and 1B, respectively.

After the list of test words are randomized into the three sub-sets described above, the five choices for each test item were randomized. The entire process was done with a table of random numbers.

In addition to the two hundred word list, twenty practice items were placed at the beginning of the test. These practice sets were selected from the pool of test items that had been rejected from the test because they did not fully meet the criteria. They appear to the subjects to be an integral part of the test, but they are not scored. These items were selected to familiarize the listener with the nature of the task and the types of words encountered on the test. Samples of the three types of test items--initial and terminal consonants, and vowels--appear on the practice list. Since proper names and slang words appear occasionally as test items or foils on test sets, they are included also in the practice sets. The complete answer sheets for Test Order 1A are presented in Appendix A.

The instructions which appear on the first page of the answer sheets are also presented at the beginning of the audio-tape. Subjects are instructed to circle one, and only one, word from each row. They are told to respond to every stimulus and to guess if necessary.

Recording

The 220-word list (practice sets included) was recorded by a male speaker native to New York City. He is trained in phonetics and experienced in this type of recording procedure.

The speaker was seated approximately six inches from an Electrovoice Model 654 microphone in a quiet room. He did not use a VU meter to monitor his voice during the recording. However, the experimenter, in another room with the tape recorder, noted any words which she judged to be extremely high or low on the tape recorder VU. At the end of the first reading, the few words that had to be re-recorded were presented again to the speaker. The recording was made on an Ampex Model AG500 recorder at a tape speed of 7-1/2 i.p.s.

The original tape recording was dubbed onto a second matched Ampex AG500 recorder. For this copy, the experimenter copied each word individually, having adjusted the reproduce level of recorder #1 and/or the record level of recorder #2 so that each word peaked at -3dB ± 1dB. A calibration tone of 1000 Hz was recorded at the beginning to peak at 0dB.

This equalized version was cut at the beginning and end of each word. Leader tape was spliced between words so that there was a six-second pause between each block of five words, and a four-second pause between words within the blocks. Eight-second pauses were placed after the last word on each page of the answer sheets. Playing time for the entire test, including instructions and practice items, is approximately seventeen minutes.

Each stimulus word is presented alone with no carrier phrase. Previous investigation of the value of a carrier phrase (Martin, Hawkins and Bailey, 1962) showed that there was no difference in the scores obtained on PB lists with and without the carrier phrase. Only eighteen of their seventy-five hypacusic and normal listeners preferred the carrier. The otherspreferred to omit the carrier because it confused them. In the present study it was decided that a carrier phrase (which would be unfiltered) not only might confuse the listeners, but would also complicate the recording procedure and increase listening time for the subjects.

After the amplitude equalization and timing were completed, two copies of this tape (Test Orders 1A and 1B) were made for presentation to hard-of-hearing and normal-hearing subjects.

The equalized and spliced tape was then copied under sixteen filter conditions. (It should be noted that according to this procedure, all test tapes are two copies away from the original tape recording.) Eight high-frequency cut-off points were selected: 250, 500, 750, 1000, 1250, 1750 and 2000 Hz. For each of these low-pass filter settings, a copy was made using one Allison Variable Filter Model 2BR, with an effective slope of 30 dB per octave.[1] Additional copies were made with each of the eight filter settings using two of the Allison filters with a combined effective slope of 60 dB per octave.

All sixteen of these tapes were used for a pilot study with eight subjects. However, only eight tapes were presented to the larger group in the main study, namely: the two low-pass filter tapes of 1000 and 1500 Hz with a 30 dB/ octave slope, plus the six filter tapes of 500, 750, 1000, 1250, 1500, and 2000 Hx with a 60 dB/ octave slope.

The unfiltered tape was presented in both the pilot study and the main study, as well as to a small unselected group of hard-of-hearing individuals.

Subjects

In addition to the eight subjects who participated in the pilot study, the 104 normal-hearing subjects were students and faculty at the City University of New York. No subject had a history of ear pathology or hearing loss. All were native speakers of American English. All were natives of the New York Metropolitan area, or in the cases of some of the older subjects, had lived there most of their adult lives.

These subjects ranged in age from 16 to 63 years.

---

[1] If the filter is set to remove frequencies above 1000 Hz, the frequency which is an octave higher, in this example, 2000 Hz, will be 30dB less intense than 1000 Hz.

The age range and mean age of subjects for each test condition are given in Table 14.

The Phoneme Differentiation Test was also administered to nine hard-of-hearing subjects. All were outpatients at the Temple University Medical School Audiology Department. The patients were selected immediately after they had received their clinical tests and were given the PD Test on the same day. All patients had received the necessary medical examinations and whatever further testing was needed for diagnostic evaluation. Not all of these patients seemed ideally suited to this type of test because of age or intellectual ability, but they were tested nonetheless, because the time was available and because the examiner wanted to find out if this type of testing was feasible with them. Though they ranged in age from 10 to 77 years, it appears that all were able to handle the task.

The Administration of the Test

The normal-hearing subjects were accepted for testing after they had passed an audiometric screening via a Maico MA-2B audiometer. The presentation level for the screening was 15 dB at: 500, 1000, 2000 and 4000 Hz.[2]

In an IAC sound-treated room, the test lists were

---

[2] Two of the subjects did not respond at 15 dB at 4000 Hz. These subjects were given the PD test but note was taken of their hearing levels (25 dB in both cases). These slight losses at a very high frequency did not appear to affect scores, especially since these frequencies were filtered out.

TABLE 14

AGE RANGE AND MEAN AGE OF SUBJECTS FOR EACH TEST CONDITION

Test Conditions Filter Setting	 Filter Slope	Number of Subjects	Age Range in Years	Mean Age in Years
500 Hz	60 dB	12	19-54	25.1
750 Hz	60 dB	12	18-54	25.6
1000 Hz	60 dB	12	19-49	23.3
1250 Hz	60 dB	12	16-35	22.6
1500 Hz	60 dB	12	19-63	28.4
2000 Hz	60 dB	12	19-46	25.6
1000 Hz	30 dB	10	19-50	24.0
1500 Hz	30 dB	10	19-51	27.6
Unfiltered		12	18-48	24.8
Total Number of Conditions		Total No. of Subjects	Range for All Cond.	Mean Age for All Cond.
9		104	16-63	25.2

presented monaurally through TDH-39 earphones with MX/41AR cushions. Playback equipment included an Ampex AG500 tape recorder, with a McIntosh 75 Amplifier and a Hewlett Packard Attenuator. The presentation level for all lists was approximately 56 dB SPL, as measured with a Bruel & Kjaer Sound Level Meter Type 2203 on the "A" scale. This level was chosen because it coincided with that used for the hard-of-hearing subjects, and it approximates the sensation level used for speech discrimination testing in most audiology clinics.[3] This would allow comparison of scores on the PB and PD tests.

Each normal-hearing subject listened to the complete PD test under only one of the filter conditions. The distribution of subjects is shown in Table 14. Twelve subjects listened to the test under each of the following filter conditions: unfiltered, low-pass filtered at 500, 750, 1000, 1250, 1500, and 2000 Hz, using a 60 dB/ octave slope. Ten subjects listened to the test under each of the following: 1000 and 1500 Hz using a 30 dB/octave slope. Half of the listeners for each of the nine filter conditions were given Test Order 1A. The other half were given Test Order 1B.

Distribution of subjects across filter conditions

---

[3] The speech level of 0 dB on the audiometer used in the present study is equal to 20 dB SPL. Thus, for example, an individual with an SRT = 20 dB would receive the Phoneme Differentiation Test, as well as other speech discrimination tests, at 76 dB SPL.
(20 dB SPL + 20 dB SRT + 36 dB SL = 76 dB SPL)

was balanced for age and listening experience. That is, for each test condition, no more than two listeners over age 35 years were tested. Those subjects who had previously participated in filtered-speech experiments (faculty and graduate students) were distributed across filter conditions so that at least two, but not more than three experienced subjects listened under any one filter condition.

Subjects were allowed to listen with the preferred ear. Those with a preference generally chose the ear used for the telephone. Those subjects without a preference were assigned the test ear by the experimenter, so that half of the subjects for any filter condition listened with the right ear and half listened with the left.

The nine hard-of-hearing subjects were given the PD test by the experimenter or one of the regular members of the clinic staff. The tests were administered in sound-treated IAC rooms, via a Maico MA-24 Diagnostic Audiometer with a Viking tape deck. The unfiltered list (either Test Order 1A or 1B) was presented monaurally via TDH-39 earphones with MX41AR cushions. Since all 200 test items were presented in one list, only one ear for each patient could be tested.[4] Selection of the ear for test was dependent upon the degree and type of hearing loss and the discrimination score on PB lists. In some cases, the poorer ear was

---

[4] In future clinical applications of the test, four lists of 50 words each would probably be used. Thus both ears could be tested.

tested, while in other cases, the better ear was used. Masking in the contralateral ear was used only when it was considered necessary. Masking was used with any subject for which presentation level of speech stimuli, minus the intra-aural attenuation of 40 dB, was above the best bone conduction threshold of the non-test ear, (Martin, 1968). All hard-of-hearing subjects received the PD test at 36 dB above their speech reception thresholds.

CHAPTER IV

RESULTS

The Normal-Hearing Listeners

The Pilot Test

A preliminary investigation was conducted to determine the filter conditions to be used in the main study. Eight City University students aged 19 to 25 years were each given the Phoneme Differentiation Test[1] under the sixteen filter conditions described in Chapter III. So that learning of the list would be minimized, the 200 test items were presented in a different random order for each of the conditions. The eight subjects listened to the sixteen tests in different random orders, half of the subjects receiving the 30 dB slope conditions in the first session, half receiving the 60 dB slope conditions first. Each subject was tested on two different days in two four-test sessions.

The lists were presented binaurally at 85 dB SPL, this level having been set before the hard-of-hearing testing was planned. Table 15 presents the average scores for each condition. The scores obtained in the preliminary tests are all higher than the scores for the same conditions in the main study (Table 18). The average difference is 5.5%. The

---

[1] An earlier version of the PD Test, which provided six choices for each test item, was used in the pilot study.

TABLE 15

PILOT TEST SCORES AVERAGED FOR EIGHT SUBJECTS

Low-Pass Filter Setting in Hz	30 dB Slope	60 dB Slope
250	27.5%	32.3%
500	49.8%	56.4%
750	74.1%	62.7%
1000	77.1%	71.3%
1250	92.4%	82.8%
1500	92.4%	88.8%
1750	95.3%	91.9%
2000	94.9%	94.0%
All Conditions Averaged	75.5%	72.5%

differences result from the fact that the pilot tests were presented at a level of intensity 29 dB higher than that used in the main study. In addition, the pilot lists were presented binaurally, while in the main study the lists were presented monaurally.

Table 15 shows that the pilot scores increase as the low-pass cut-off frequency is raised. For the full-test scores there is only one reversal, 1750 Hz and 2000 Hz with the 30 dB slope. The filter setting of 1750 Hz was one of those eliminated from the main study.

All of the scores are above chance level (16.7% for a six-choice test). Inspection of Tables 16 and 17, which give the scores for consonants and vowels, respectively, shows that the consonant scores for the 250 Hz conditions are substantially higher than the vowel scores. That is, vowels are more adversely affected by elimination of frequencies between 250 and 500 Hz than are the consonants. While it would be interesting to verify this observation, the 250 Hz conditions were not used in the main study because it was difficult to listen under these conditions when there was a lower-level monaural presentation. Two inexperienced subjects given the 250 Hz test at the beginning of the main study were not able to complete it. A subject with experience in listening to filtered speech was able to complete the test, scoring 32.5%. Her scores of 23.0% for vowels and 36.4% for consonants lend support to the earlier observation that vowels are more adversely affected by elimination of

TABLE 16

PERCENT CONSONANTS CORRECT AVERAGED FOR EIGHT PILOT TEST SUBJECTS

Low-Pass Filter Setting in HZ	All Consonants		Initial Consonants		Terminal Consonants	
	30 dB Slope	60 dB Slope	30 dB Slope	60 dB Slope	30 dB Slope	60 dB Slope
250	32.8%	36.6%	29.5%	36.8%	36.1%	36.5%
500	46.6%	53.4%	52.6%	53.7%	40.6%	53.0%
750	68.5%	58.3%	71.0%	67.0%	65.9%	49.5%
1000	73.4%	65.3%	74.9%	68.6%	71.9%	62.0%
1250	89.7%	78.5%	91.3%	81.0%	88.0%	76.0%
1500	90.9%	86.2%	92.3%	83.1%	89.4%	87.3%
1750	93.1%	89.4%	94.5%	89.6%	91.6%	89.2%
2000	92.0%	91.4%	94.6%	91.5%	89.4%	91.3%
All Conditions	73.4%	69.9%	75.1%	71.4%	71.6%	68.1%

TABLE 17

PERCENT CORRECT FOR ALL VOWELS AND DIPHTHONGS AVERAGED FOR EIGHT PILOT TEST SUBJECTS

Low-Pass Filter Setting in Hz	30 dB Slope	60 dB Slope
250	17.1	22.1
500	53.5	63.3
750	85.6	67.7
1000	85.0	83.3
1250	97.9	91.3
1500	96.5	98.3
1750	99.9	97.7
2000	100.0	100.0
All Conditions Averaged	79.4	78.0

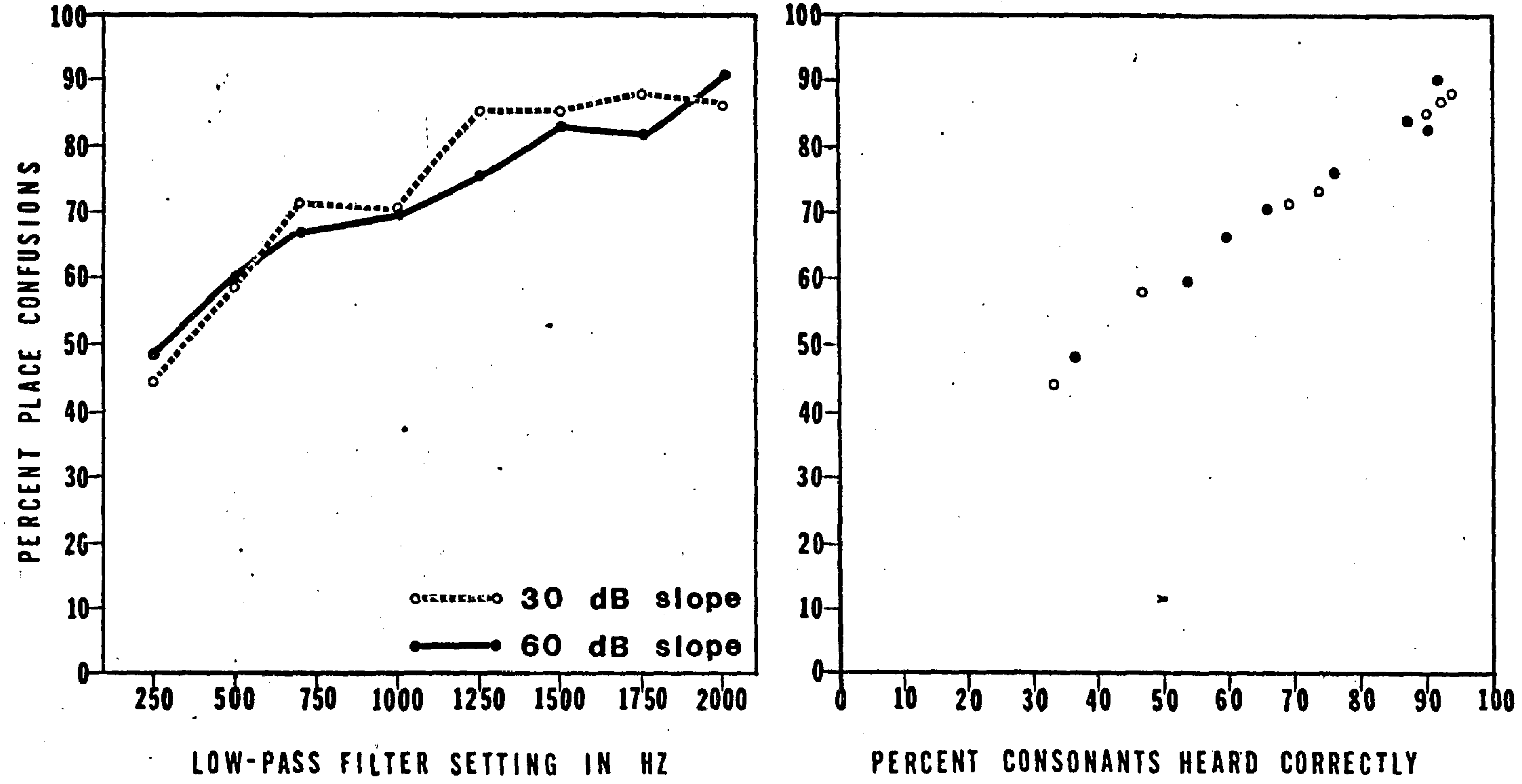
very low frequencies than are consonants.

It is possible that not only one but a group of similarly experienced listeners could have completed the tests with the 250 Hz setting. However, use only of experienced listeners for these test conditions would have necessitated use of experienced listeners for all the test conditions. Since such a large number of these subjects was not available, further investigation of the 250 Hz conditions was postponed for future study.

Comparison of the scores for the two filter slopes used in the pilot study (Table 15) shows that the scores obtained when the filter slope is 60 dB are generally poorer than those obtained when the slope is 30 dB. However, there are two reversals from the expected direction. For the 250 and 500 Hz settings, the scores were better with the 60 dB slope. Other reversals are found in the vowel and consonant data at other frequencies as well (Tables 16 and 17). For the six cut-off frequencies later used in the main study--500, 750, 1000, 1250, 1500, and 2000 Hz--the average difference between scores is 4.1% in favor of the 30 dB slope.

In general, it appears that although scores are slightly lower for the sharper slope, the types of errors made are very similar. Figure 2.a compares the percentage of place confusions for the two slopes plotted against the various filter settings. Figure 2.b compares the percentage of place confusions plotted against the percentage of consonants heard correctly for the two slopes. The proportion of

Figure 2. Proportion of pilot test consonant place confusions for the 30 dB and 60dB slopes plotted as a function of filter setting (2.a) and as a function of consonant intelligibility (2.b)


PERCENT PLACE CONFUSIONS
100
90
80
70
60
50
40
30
20
10
0
250
500
750
1000
1250
1500
1750
2000
LOW-PASS FILTER SETTING IN HZ
30 dB slope
60 dB slope
PERCENT CONSONANTS HEARD CORRECTLY
0
10
20
30
40
50
60
70
80
90
100

a

b

total consonant errors which are place confusions tends to increase as the filter setting is raised. The same tendency appears when the proportion of place confusions is plotted against the percentage of consonants heard correctly. Comparison of figures 2.a and 2.b shows that the slight difference in the proportion of place confusions in favor of the 30 dB slope is apparently a reflection of the slightly higher overall scores obtained. For both slopes, the proportion of place errors increases as a function of increasing intelligibility. Thus, the two slopes appear to yield slight differences in scores, but no differences in error type.

Since little additional information was obtained by using two slopes, the 60 dB slope was used in the main study. However, two filter settings, those at 1000 Hz and 1500 Hz, were tested, using both slopes, to check for any significant differences which might emerge from testing a much larger sample.

The Main Study

The major part of this chapter consists of an analysis of the 20,800 responses of 104 normal-hearing subjects under the various conditions of frequency distortion.

Two test orders were used. Half of the subjects for each condition were given Test Order 1-A: (1) initial consonant test items; (2) vowel and diphthong test items; (3) terminal consonant test items. The remaining subjects for each condition were given Test Order 1-B: (1) vowel and diphthong test items; (2) terminal consonant test items; (3) initial consonant test items. The difference in total errors between the test orders is approximately 0.2%.

Group 1-A, which heard the initial consonant test items first, made 1.5% more errors on these items than the group which heard them last. This is probably a result of their learning the task, which was in this case, identifying filtered words. A similar warm-up effect appears for the vowel test items. Group 1-B, which heard the vowel items first, made 2.7% more errors on these items than did group 1-A, which heard them in the middle of the list.

The 1.2% more errors for terminal consonants by group 1-A, which heard these last in the list, might be a result of greater fatigue than for group 1-B, which heard the terminal consonants in the middle of the list.

In general, the differences between test orders are very slight, and it is probable that in future applications of the Phoneme Differentiation Test, using shorter lists, the differences would not appear. The lack of any large warm-up effects can probably be attributed to the use of practice items immediately preceding the test. The results for the two test orders are averaged in all of the data presented below.

The average scores for all test conditions are shown in Table 18. The scores for the unfiltered condition are all better than 95.0% for the sub-tests, averaging 98.0% for the full list. It can therefore be assumed that the errors found for all other listening conditions are primarily due to the

TABLE 18

PERCENT CORRECT RESPONSES FOR EACH MAIN TEST FILTER CONDITION

Test Condition		Initial Consonants	Terminal Consonants	Vowels	Diphthongs	Full Test
Filter Setting	Filter Slope					
500 Hz	60 dB	52.8	44.2	42.9	32.5	46.9
750 Hz	60 dB	58.6	51.0	60.0	63.3	57.4
1000 Hz	60 dB	61.9	56.1	68.3	63.8	61.9
1250 Hz	60 dB	80.2	72.4	82.9	90.0	79.7
1500 Hz	60 dB	83.0	76.4	98.1	99.6	86.0
2000 Hz	60 dB	89.6	91.1	98.1	99.2	92.7
1000 Hz	30 dB	65.6	61.4	76.5	81.5	68.3
1500 Hz	30 dB	87.1	81.2	96.3	97.5	88.5
Unfiltered		98.4	95.5	99.4	100.0	98.0
All Conditions Averaged		75.3	69.9	80.0	80.5	

frequency distortion, and not to such factors as: the ideolect or dialect of the speaker, the recording technique, the instrumentation, the word list, the nature of the task, or the level of intensity at which the test lists were presented.

Comparison of the data for the two filter slopes shows consistent differences in favor of the 30 dB slope, except for the two reversals at 1500 Hz for vowels and diphthongs. The average difference is 4.5% in favor of the less steep slope. The largest difference is the 17.7% increase in score for diphthongs when the 30 dB slope is used at 1000 Hz. This would suggest that for correct perception of the diphthongs tested much of the needed acoustic information is contained in the frequencies just above 1000 Hz.

When the scores for all filter settings are averaged (Table 18), it appears that vowels are more intelligible than consonants, and furthermore, that initial consonants are more intelligible than terminal consonants. The advantage of initial over terminal word position for consonants is found in eight of the nine test conditions. The one reversal (at 2000 Hz) is less than 2.0%. The terminal consonants are not as intelligible as the initial consonants even in the unfiltered listening condition. Vowels and diphthongs are more intelligible than consonants for all conditions except 500 Hz. A similar reversal was found in the pilot data with the 250 Hz setting, but not with the 500 Hz setting. This suggests that the reverse order intelligibility of consonants over vowels occurs when the overall score is low enough--46.9% at

500 Hz in the main study, and 32.3% at 250 Hz in the pilot study. It may also suggest that the lower presentation level in the present study caused more frequencies to fall below sensation level, thereby simulating a lower cut-off frequency.

Figure 3 shows the regular increase in score with increase in cut-off frequency for all conditions tested. In addition, it shows the increase in overall score when the filter slope is less sharp. The average difference of 4.5% between slopes for the two filter settings is similar to the average difference of 4.1% between slopes in the pilot data. As for type of error, no differences were found between slopes either in the pilot project or in the main study data when it was subjected to preliminary inspection. Therefore, the 30 dB slope conditions were not considered in the more detailed analyses which follow.

The articulation gain functions[2] for vowels and consonants are shown in Figures 4 and 5, respectively. The apparent linearity of the full-test curve above 1500 Hz results from the combined effects of the sharp increase in intelligibility for consonants and a leveling of the rather high intelligibility for vowels and diphthongs. Thus, over 1500 Hz, there is very little increase in intelligibility for vowels and diphthongs (less than 2.0%), while for consonants there is much information to be gained from frequencies above 1500 Hz (approximately 17.0%).

---

[2] Plots of intelligibility scores for a set of listening conditions which increase in intensity or spectrum.

Figure 3. Percent correct responses for each main test filter condition. For each 60 dB filter condition, 2400 stimuli were presented. For each 30 dB filter condition, 2000 stimuli were presented.

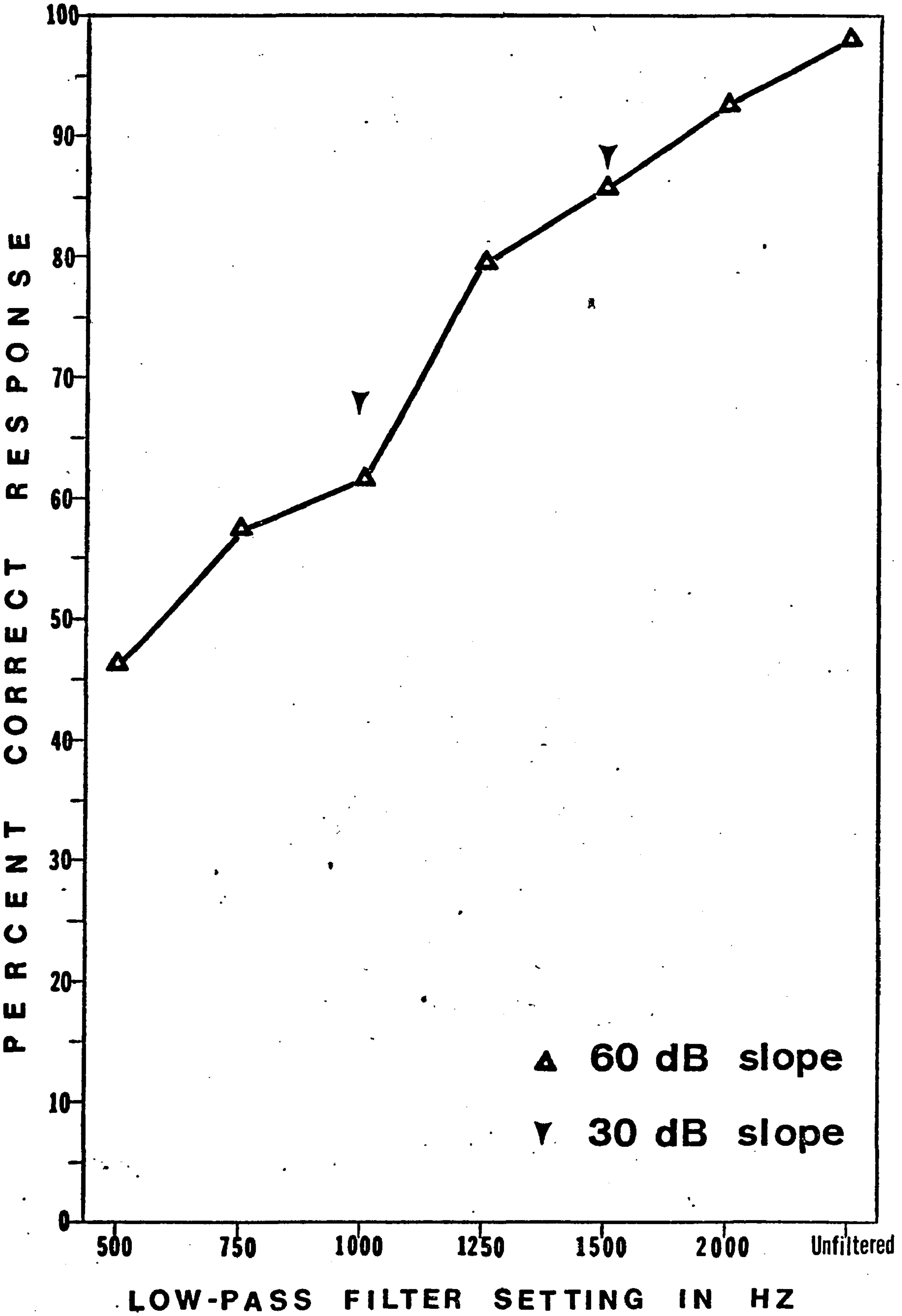
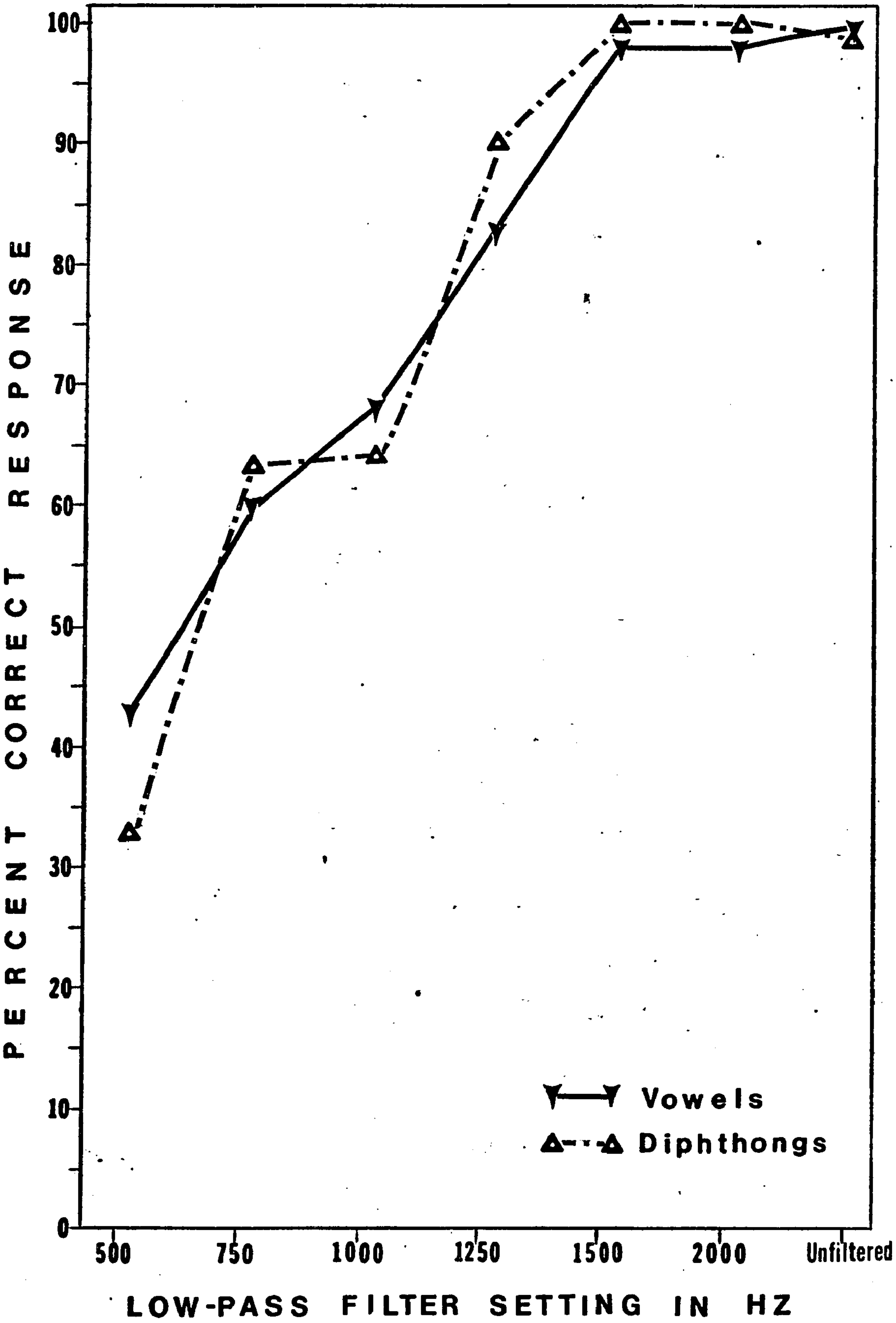
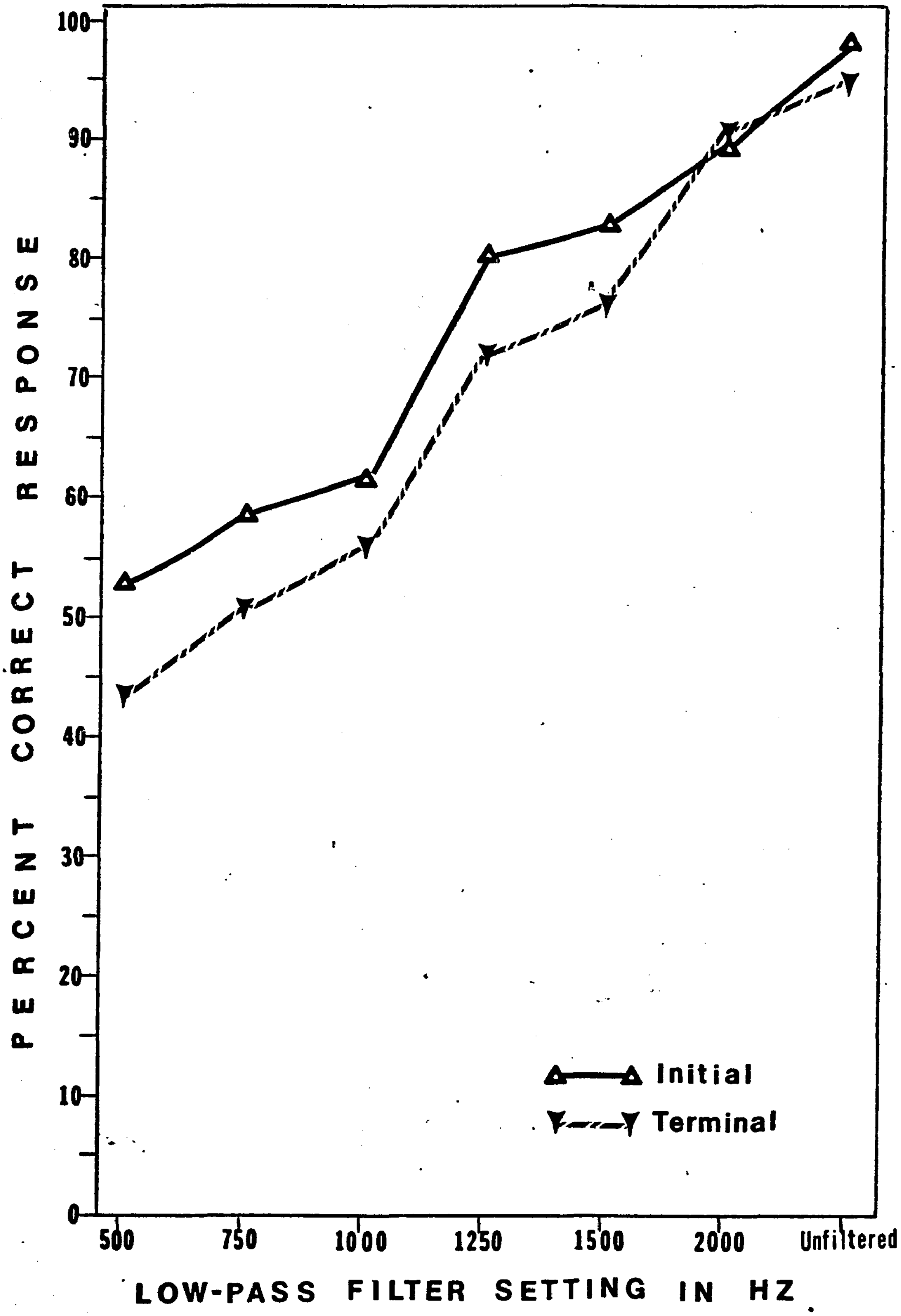

PERCENT CORRECT RESPONSE
100
90
80
70
60
50
40
30
20
10
0
500
750
1000
1250
1500
2000
Unfiltered
LOW-PASS FILTER SETTING IN HZ
▲ 60 dB slope
▼ 30 dB slope

Figure 4. Percent correct responses for vowels and diphthongs. For each filter condition, the number of vowel stimuli was 480; the number of diphthong stimuli was 240.



PERCENT CORRECT RESPONSE
100
90
80
70
60
50
40
30
20
10
0
500
750
1000
1250
1500
2000
Unfiltered
LOW-PASS FILTER SETTING IN HZ
Vowels
Diphthongs

Figure 5. Percent correct responses for initial and terminal consonants. For each filter condition, 1056 initial consonant stimuli were presented; 624 terminal consonants were presented.


PERCENT CORRECT RESPONSE
100
90
80
70
60
50
40
30
20
10
0
500
750
1000
1250
1500
2000
Unfiltered
LOW-PASS FILTER SETTING IN HZ
Initial
Terminal

The Consonants.- The relative contributions of the features place, openness, nasality, and voicing to the error rates for consonants are illustrated in Figures 6 and 7. For each test condition, all errors were analyzed according to the feature confused. The percentages shown are the number of each feature confusion divided by the total number of errors for that condition. The number of errors made by all twelve subjects for each condition are shown below the low-pass frequency setting in the figures.

The most apparent trend is an increase in the proportion of place confusions accompanied by a corresponding decrease in openness, nasality, and voicing confusions, as the listening condition improves. This is true for consonants in both initial and terminal-word positions. Scatter and trend reversals are found for the unfiltered condition, probably because a very small number of errors are being analyzed. Since the error rate for all phoneme classes, except terminal consonants, was between 0.0% and 2.0% when the list was not filtered, this condition was not considered in the more detailed analyses.

Figures 6 and 7 also show that the turning point for the increase of place confusions and decrease of other types of error occurs at the 1000 Hz setting for initial consonants and at the 1250 Hz setting for terminal consonants. Hence, the data suggest that when frequencies above 1000 or 1250 Hz are filtered out, one can predict that place confusions will be in the clear majority, and, in addition, that openness

Figure 6. Proportion of initial consonant errors involving each of four features for each filter condition. The total number of initial consonants presented was 9152. Number of initial consonant errors is shown below each filter setting. The lower row of numbers is an upper bound on the expected standard error for the measured errors at that filter setting. The upper bound corresponds to the standard error for p.=0.5.

* Less than 2% error.

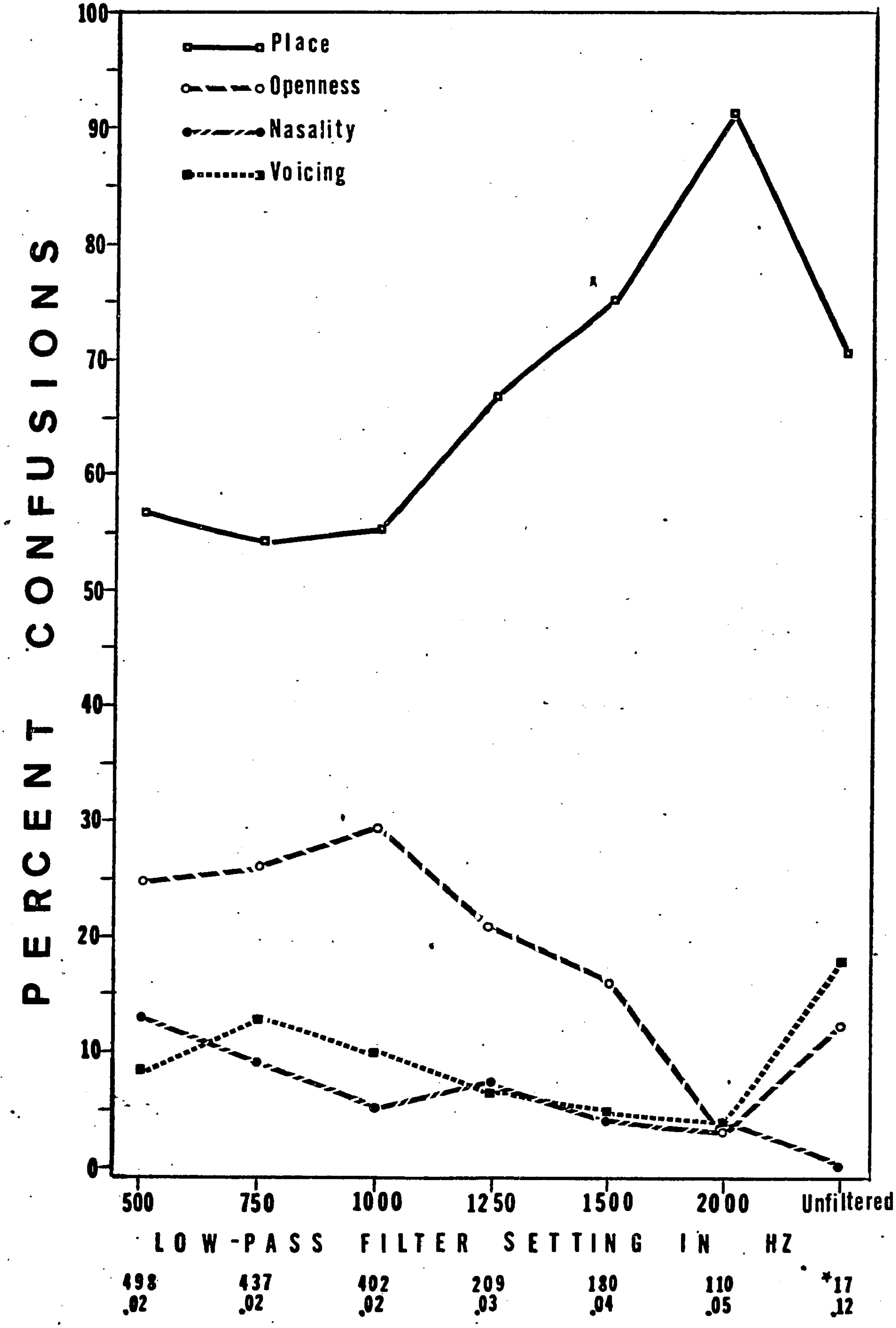
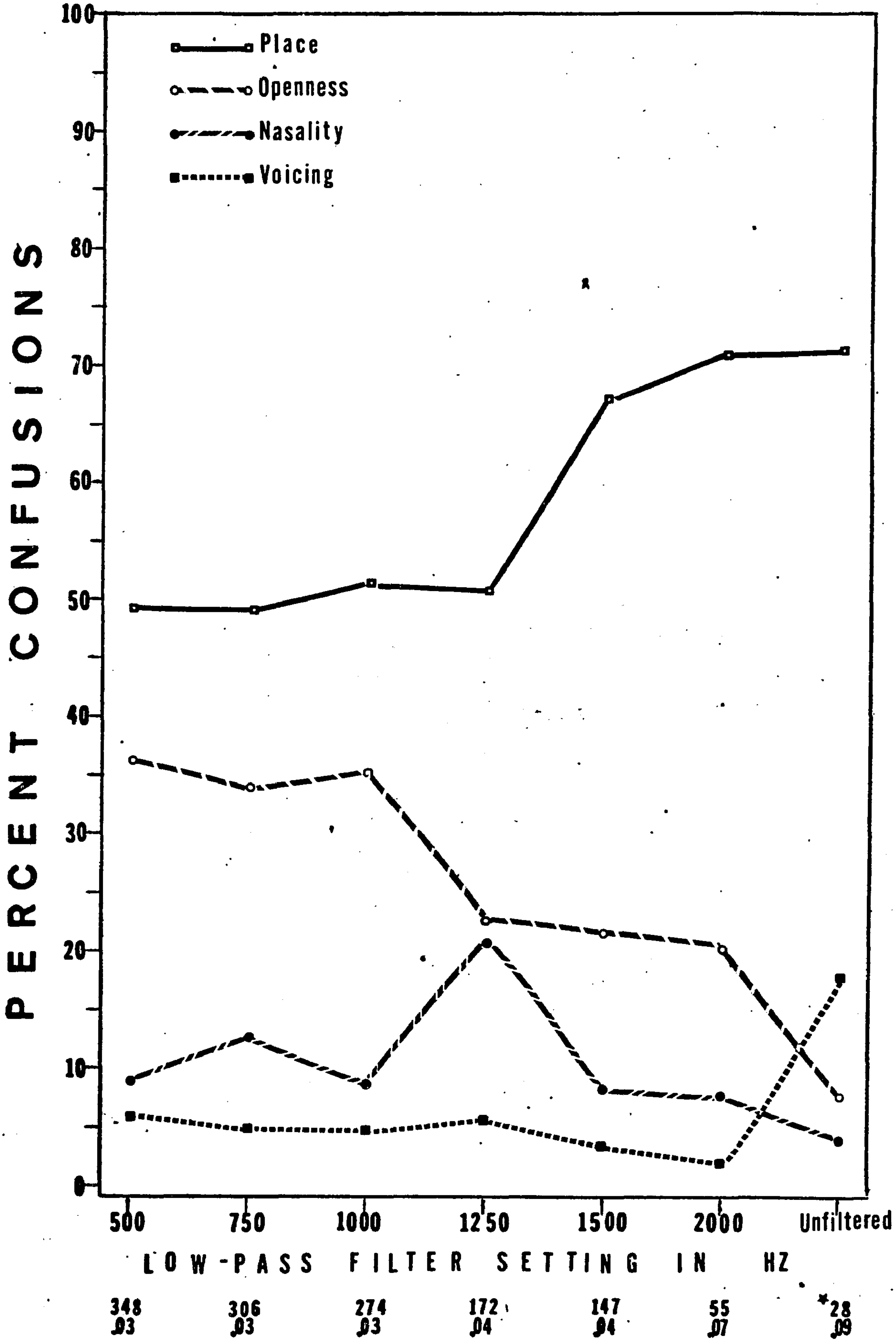

Place
Openness
Nasality
Voicing
PERCENT CONFUSIONS
100
90
80
70
60
50
40
30
20
10
0
500
750
1000
1250
1500
2000
Unfiltered
LOW-PASS FILTER SETTING IN HZ
498
.02
437
.02
402
.02
209
.03
180
.04
110
.05
*17
.12

Figure 7. Proportion of terminal consonant errors involving each of four features for each filter condition. The total number of terminal consonants presented was 5468. Number of terminal consonant errors is shown below each filter setting. The lower row of numbers is an upper bound on the expected standard error for the measured errors at that filter setting. The upper bound corresponds to the standard error for p.=0.5.

* Less than 2% error.


Place
Openness
Nasality
Voicing
PERCENT CONFUSIONS
100
90
80
70
60
50
40
30
20
10
0
500
750
1000
1250
1500
2000
Unfiltered
LOW-PASS FILTER SETTING IN HZ
348
.03
306
.03
274
.03
172
.04
147
.04
55
.07
*28
.09

confusions will be more numerous than either nasality or voicing confusions. The relative proportions are the same for the lower low-pass settings. That is, the proportions of place, openness, and nasality or voicing confusions decrease, in that order, but the differences are less sharp than for the higher filter settings.

Comparison of the data for initial- and terminal-word positions shows that the proportions of the individual feature confusions differ according to word position. Consonants in the initial position show a higher proportion of place and voicing confusions, and a lower proportion of openness confusions for all but one filter condition. The proportion of nasality confusions for initial consonants is lower than that for terminal consonants, except for the 500 Hz filter setting. As illustrated for the pilot data, the generally higher proportion of place confusions for initial consonants may be a reflection of their generally higher scores. For all consonants, the general tendency, as scores improve, is an increase in place and a decrease in all other types of confusions.

The lower proportion of voicing errors for the terminal consonants, whether plotted as a function of filter setting or score, may be explained by the extra cue of longer vowel duration preceding voiced terminal consonants. This point will be discussed in greater detail as confusions are analyzed according to consonant class.

Figures 8 and 9 show the analysis of feature confusions for voiced and voiceless consonants, respectively. For the voiced consonants there does not appear to be any great difference between initial and terminal consonants. However, Figure 9 shows that for voiceless consonants there are substantial differences in the proportion of voicing errors as a function of word position. Listeners are more likely to perceive an initial voiceless consonant as voiced than they are to perceive an initial voiced consonant as voiceless. The voiceless consonant friction or explosion noise is generally of a rather high frequency (Heinz and Stevens, 1961; Liberman _et al._, 1952), and its removal by low-pass filtering causes the stimulus to be heard as the corresponding voiced consonant based on the openness and place information found in the remaining formant structure. The removal of the consonant noise in the terminal position would not be as detrimental to the perception of voicing because of vowel duration cue. The data suggest that the normal-hearing listeners utilize this vowel-duration difference for the consonants in the terminal word-position when hearing speech under distorted listening conditions. Although no such advantage is found for the terminal consonants when the speech was not distorted, this might result from the small number of errors compared.

The rather obvious difference in the proportion of nasality errors between voiced and voiceless consonants is not surprising in view of the fact that any confusion of a

Figure 8. Analysis of feature confusions for initial (8a) and terminal (8b) voiced consonants. Number of errors is shown below each filter setting. The lower row of numbers is an upper bound on the expected standard error for the measured errors at that filter setting. The upper bound corresponds to the standard error for p.=0.5.

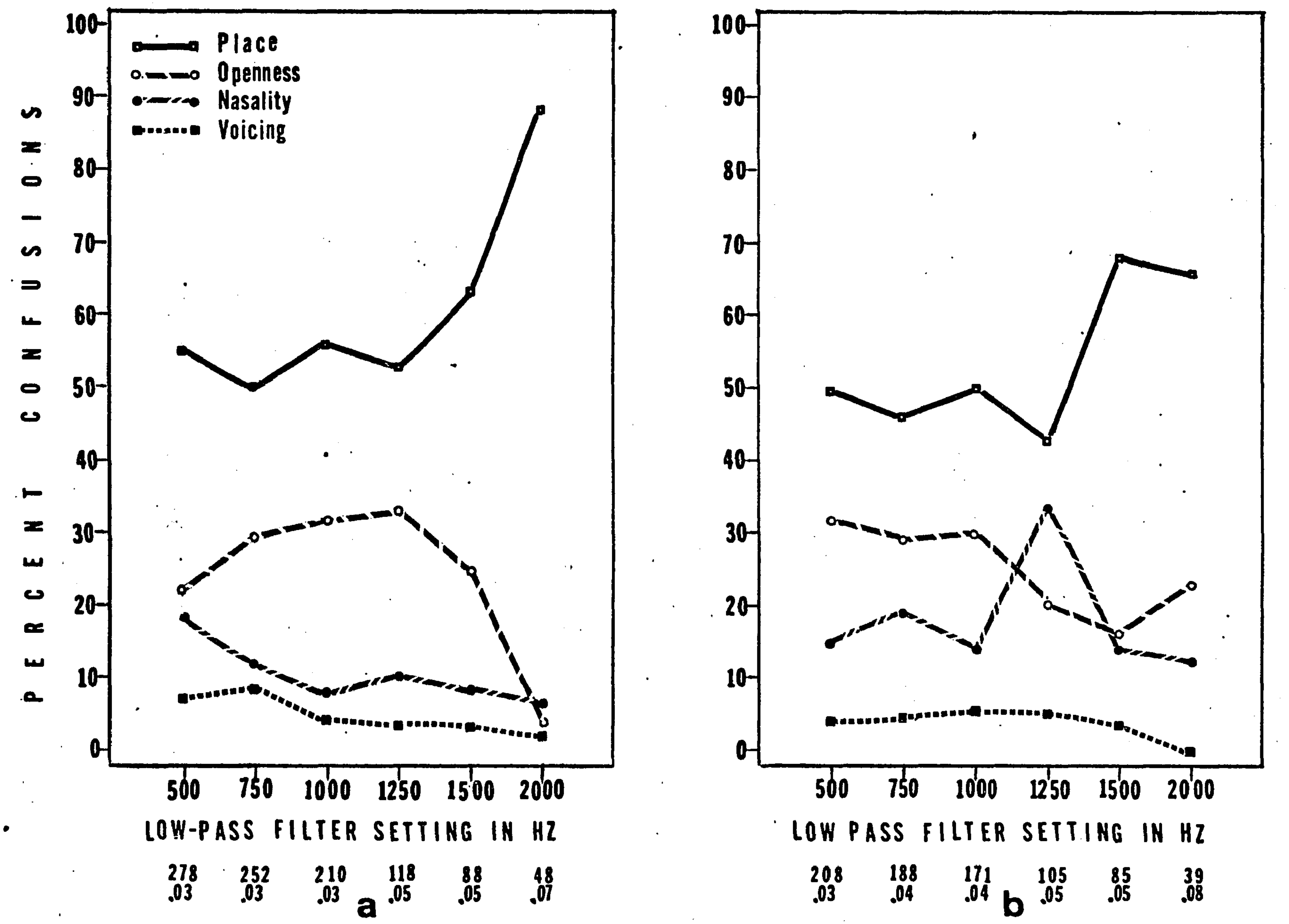
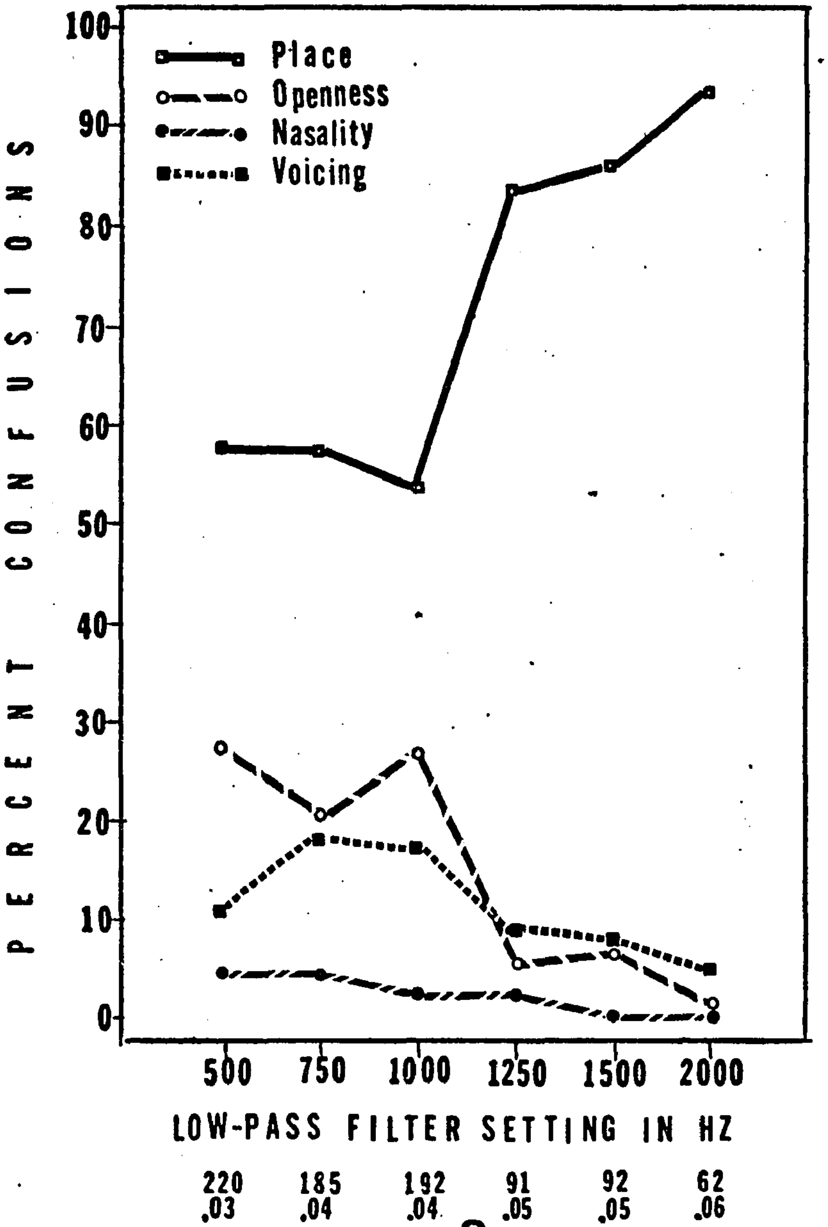
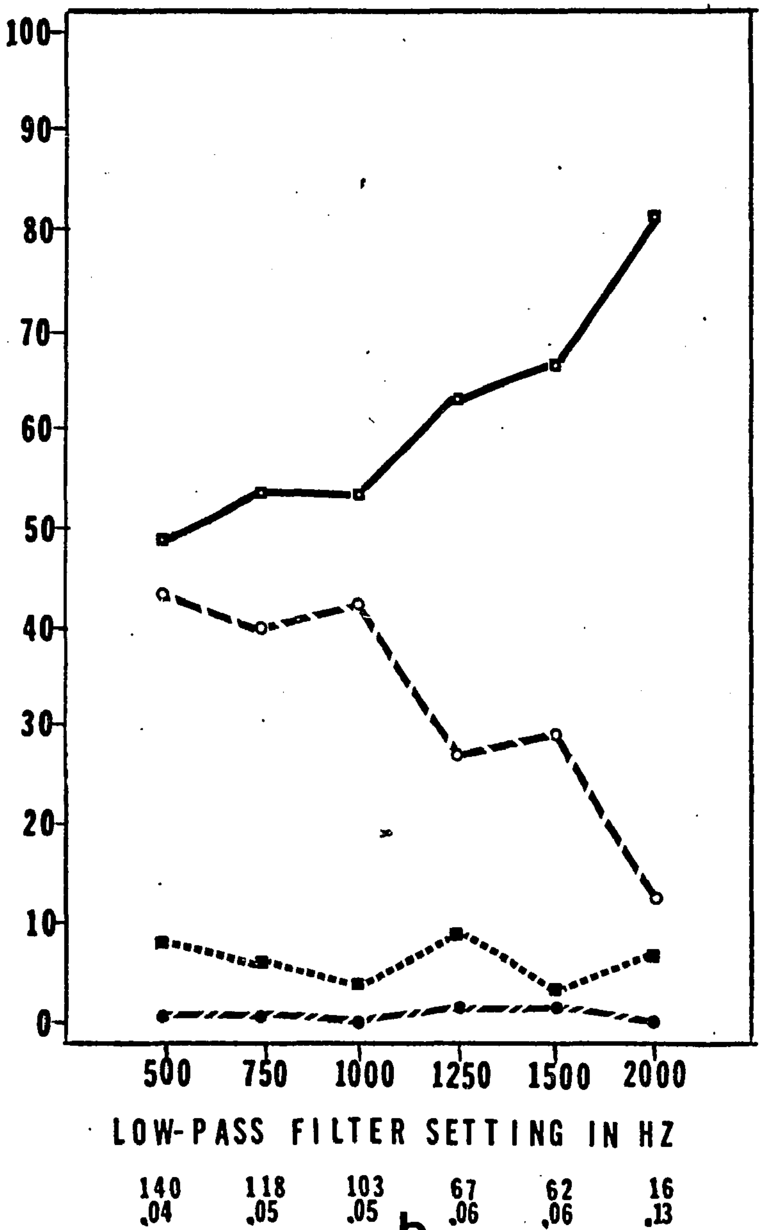

Place
Openness
Nasality
Voicing
PERCENT CONFUSIONS
100
90
80
70
60
50
40
30
20
10
0
500
750
1000
1250
1500
2000
LOW-PASS FILTER SETTING IN HZ
278 .03
252 .03
210 .03
118 .05
88 .05
48 .07
a
LOW PASS FILTER SETTING IN HZ
208 .03
188 .04
171 .04
105 .05
85 .05
39 .08
b

Figure 9. Analysis of feature confusions for initial (9a) and terminal (9b) voiceless consonants. Number of errors is shown below each filter setting. The lower row of numbers is an upper bound on the expected standard error for the measured errors at that filter setting. The upper bound corresponds to the standard error for p.=0.5.



100
90
80
70
60
50
40
30
20
10
0
PERCENT CONFUSIONS
Place
Openness
Nasality
Voicing
500
750
1000
1250
1500
2000
LOW-PASS FILTER SETTING IN HZ
220
.03
185
.04
192
.04
91
.05
92
.05
62
.06

a



100
90
80
70
60
50
40
30
20
10
0
500
750
1000
1250
1500
2000
LOW-PASS FILTER SETTING IN HZ
140
.04
118
.05
103
.05
67
.06
62
.06
16
.13

b

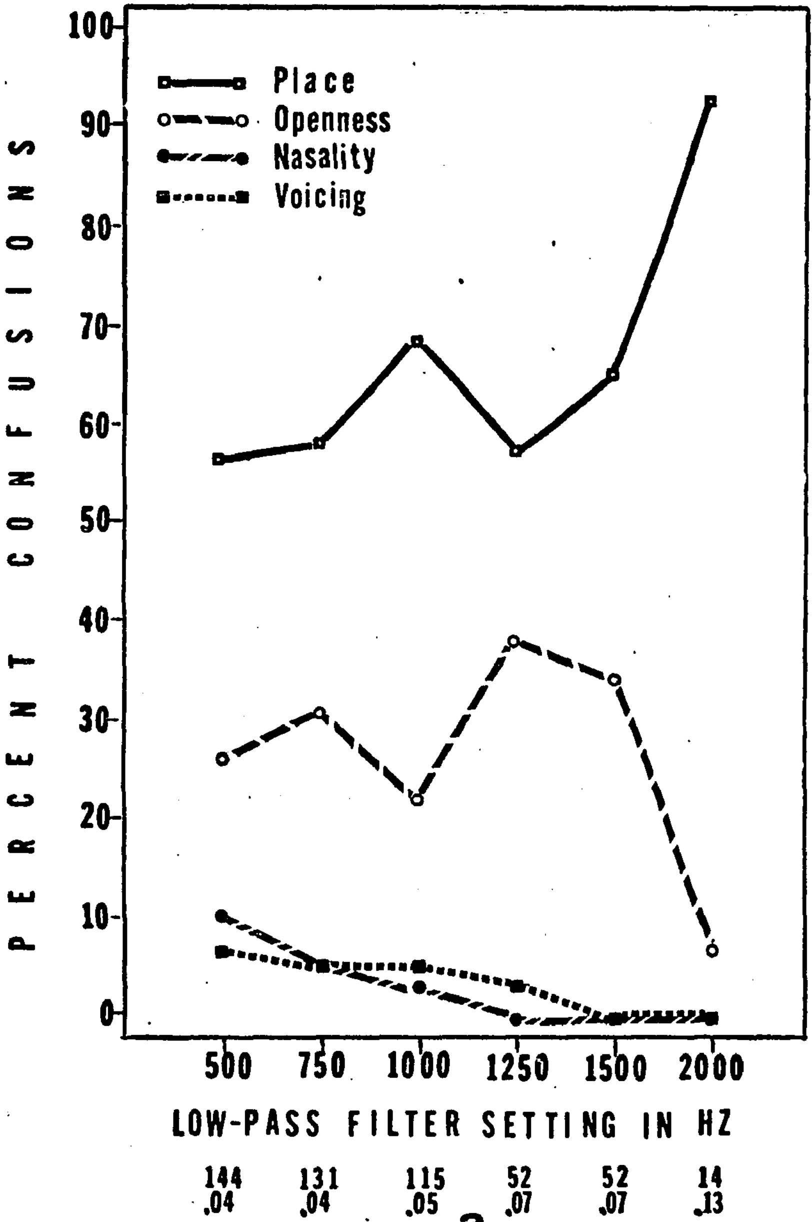
voiceless consonant with a nasal consonant would involve two features. All nasality contrasts are necessarily voicing contrasts as well. It is interesting to note that there are very few voiceless-nasal confusions in the initial position and none in the terminal position. Since one is not likely to make a voiced-voiceless confusion in the terminal position, one would be even less likely to make a voiceless-nasal confusion.

Comparison of Figures 8 and 9 also shows that there are fewer openness confusions for terminal voiced consonants than for terminal voiceless consonants. Under low-pass filtering, the openness cue is probably the first formant transition (Liberman _et al_., 1956). This transition is less likely to be audible in the voiceless consonant than in the voiced. However, this would not explain the reversal of this factor for the initial consonants. There are more openness errors for voiced than for voiceless consonants in the initial position.

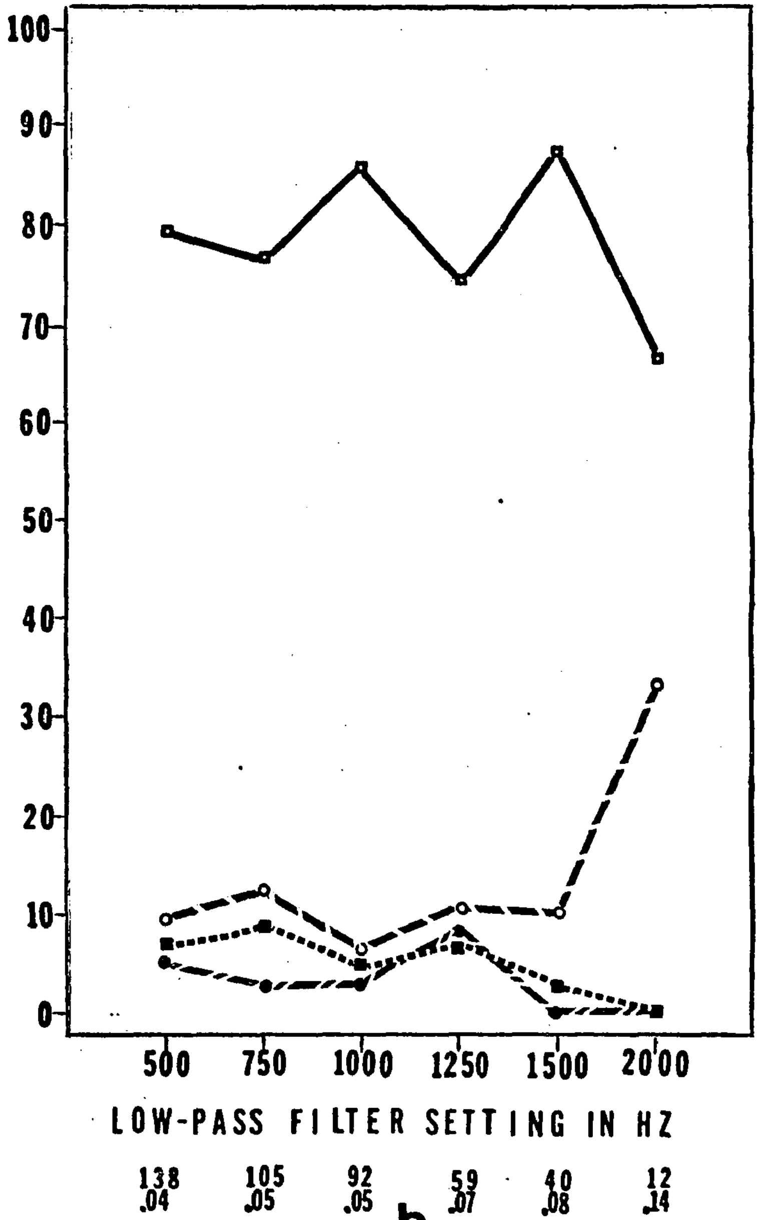
It will also be noted that in general there are more place confusions for voiceless than voiced consonants. Place cues for voiceless consonants are found in the burst friction and in the noisy second formant transition. For these stimuli, the burst and friction are probably filtered out and the noisy second formant transition may not be as perceptible as the stronger harmonic second formant transition of the voiced consonants.

Figure 10 shows the feature analysis for the

Figure 10. Analysis of feature confusions for initial (10a) and terminal (10b) plosive consonants. Number of errors is shown below each filter setting. The lower row of numbers is an upper bound on the expected standard error for the measured errors at that filter setting. The upper bound corresponds to the standard error for p.=0.5.


PERCENT CONFUSIONS
100
90
80
70
60
50
40
30
20
10
0
Place
Openness
Nasality
Voicing
500
750
1000
1250
1500
2000
LOW-PASS FILTER SETTING IN HZ
144
.04
131
.04
115
.05
52
.07
52
.07
14
.13

a


100
90
80
70
60
50
40
30
20
10
0
500
750
1000
1250
1500
2000
LOW-PASS FILTER SETTING IN HZ
138
.04
105
.05
92
.05
59
.07
40
.08
12
.14

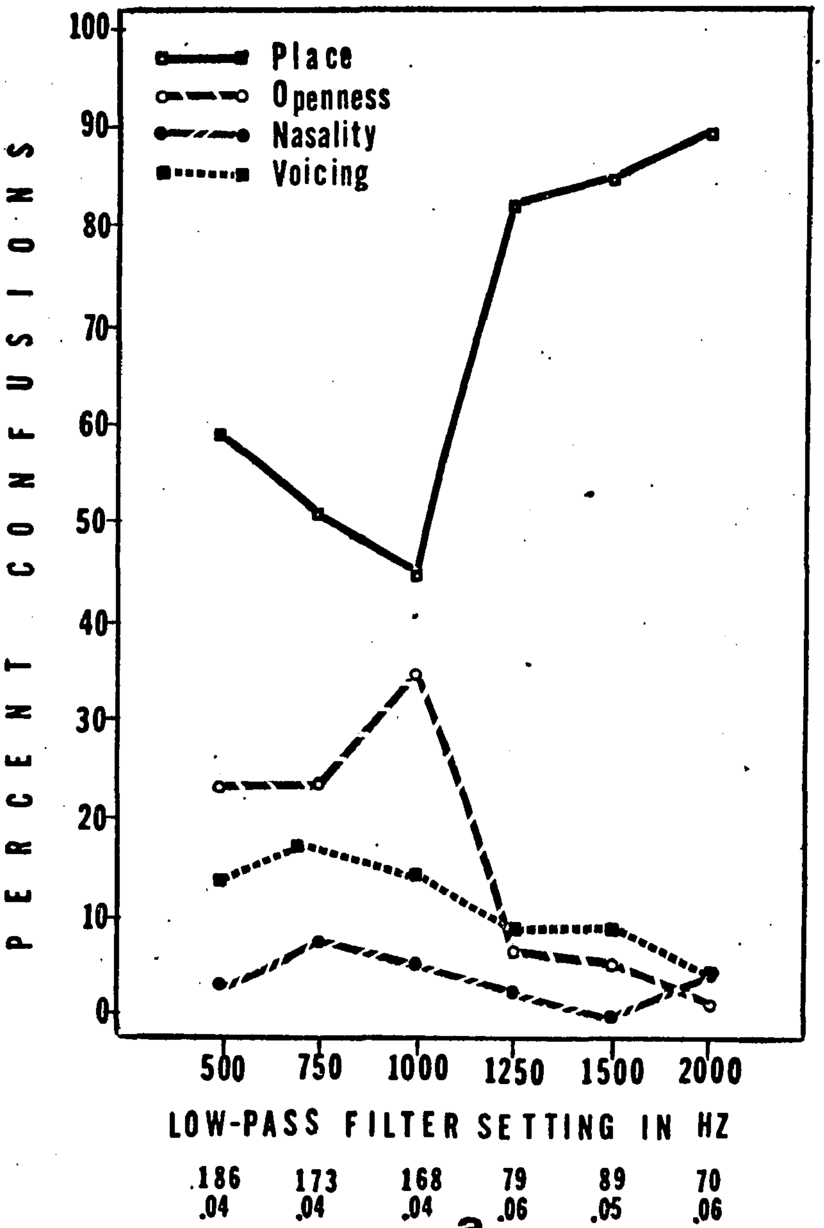
b

plosive consonants. The most striking differences between initial and terminal plosives are the greater proportion of openness confusions and the lower proportion of place confusions for the initial plosives. Listeners heard initial plosives as fricatives more frequently than they heard terminal plosives as fricatives. Figure 11 shows the opposite trend for the fricatives. Listeners heard terminal fricatives as plosives more frequently than they heard initial fricatives as plosives. The terminal fricatives are the only class of phonemes to show more openness than place confusions. This is found in the lowest three filter settings. The filtering out of frequencies above 1000 Hz would remove the fricative noises, causing the terminal fricatives to be heard as unreleased plosives. Since the removal of the consonant noise in the terminal position causes more plosive judgements, listeners respond correctly when the stimulus is a plosive, but incorrectly when the stimulus is a fricative. This would explain the higher proportion of openness confusions for fricatives than for plosives in the terminal position.

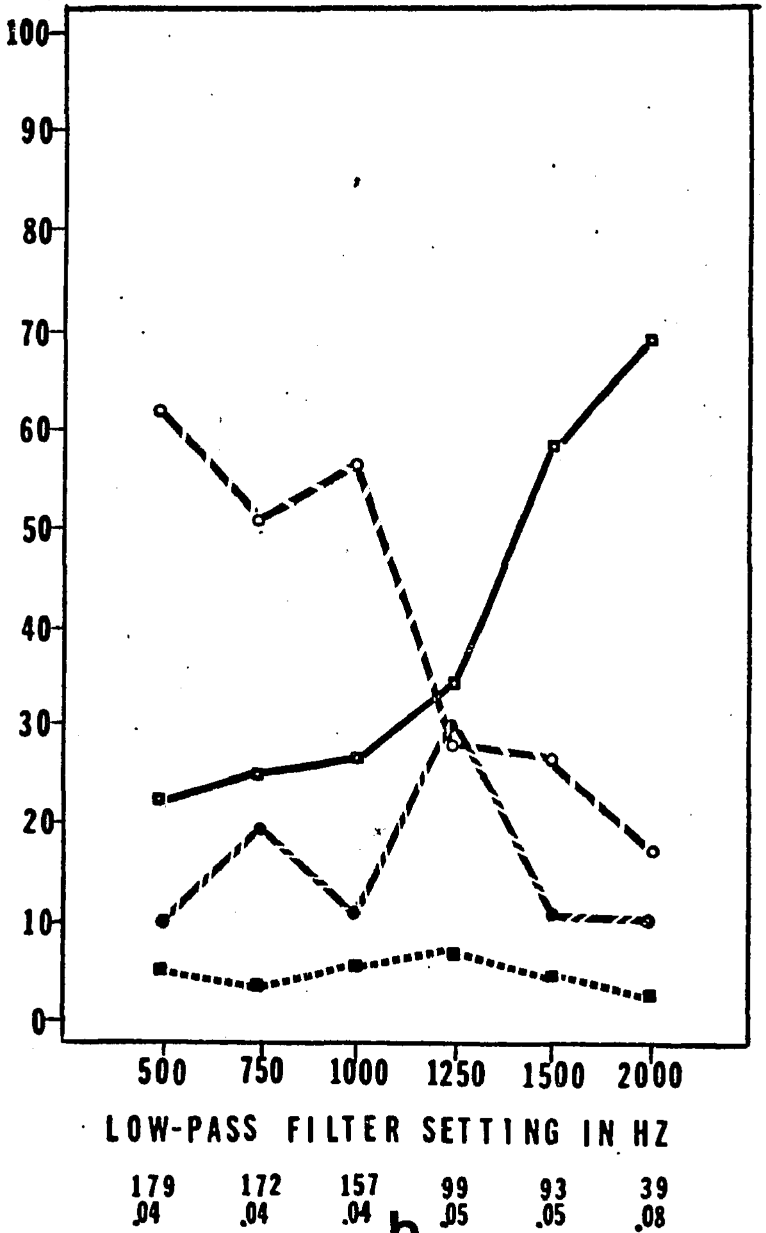
In the initial position, plosives are always released, and if the bursts were transmitted, would provide less confusion with other classes of consonants.

Analysis of the fricative confusions in Figure 11 also shows that the greater proportion of voicing errors for the initial position when all consonants are combined (as in Figure 6) is primarily a matter of voicing errors for the

Figure 11. Analysis of feature confusions for initial (11a) and terminal (11b) fricative consonants. Number of errors is shown below each filter setting. The lower row of numbers is an upper bound on the expected standard error for the measured errors at that filter setting. The upper bound corresponds to the standard error for p.=0.5.


PERCENT CONFUSIONS
100
90
80
70
60
50
40
30
20
10
0
Place
Openness
Nasality
Voicing
500
750
1000
1250
1500
2000
LOW-PASS FILTER SETTING IN HZ
186 .04
173 .04
168 .04
79 .06
89 .05
70 .06

a


100
90
80
70
60
50
40
30
20
10
0
500
750
1000
1250
1500
2000
LOW-PASS FILTER SETTING IN HZ
179 .04
172 .04
157 .04
99 .05
93 .05
39 .08

b

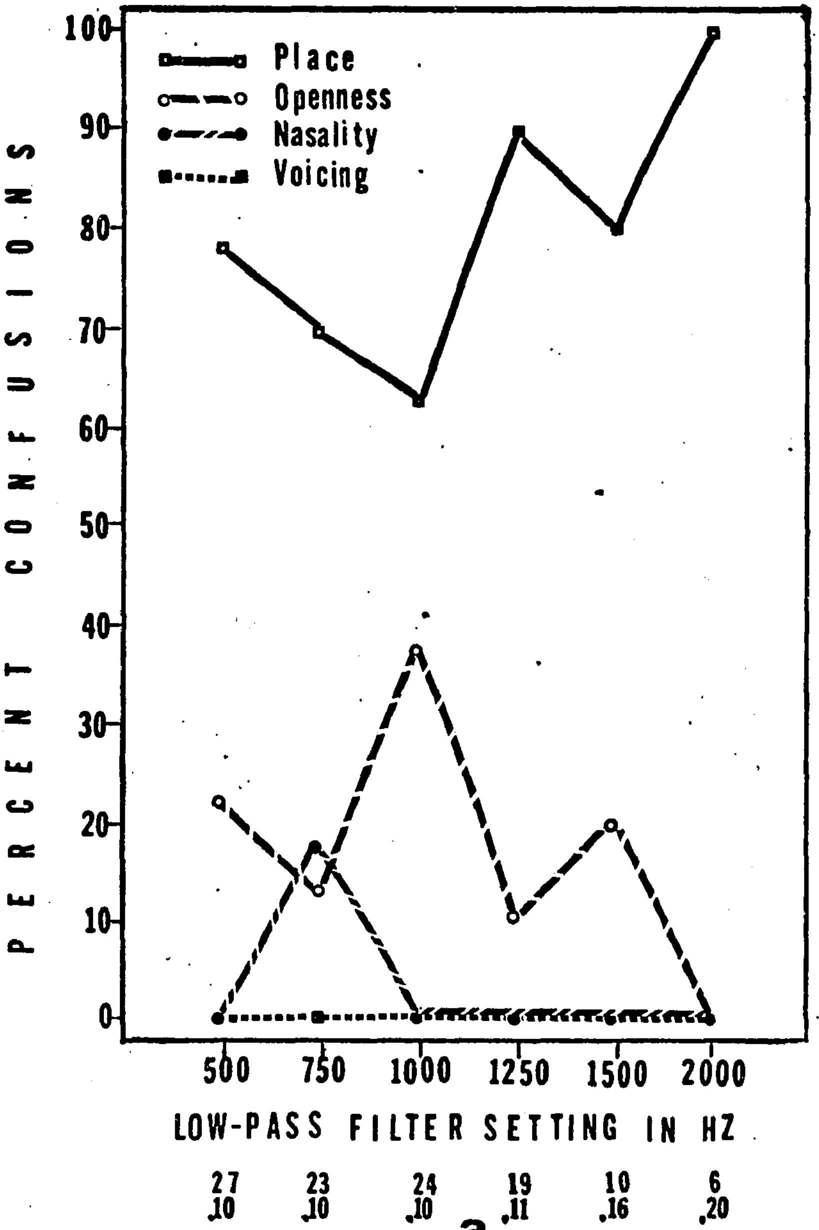
fricatives. Further, there are far fewer voicing errors for terminal fricatives than for initial fricatives. This trend is not found for the plosives. In fact, there are some minor reversals. Hence, the vowel duration cue for terminal consonant voicing operates more clearly for fricatives than for plosives.

Figure 12 shows the feature confusions for nasal consonants, The major difference between the nasals and all other classes of consonants is the complete absence of voicing confusions. As explained above, this type of confusion would involve two features. The fact that these multi-feature confusions were not found lends support to the theory that phonemes are perceived as combinations of features, each of which may be perceived independently.

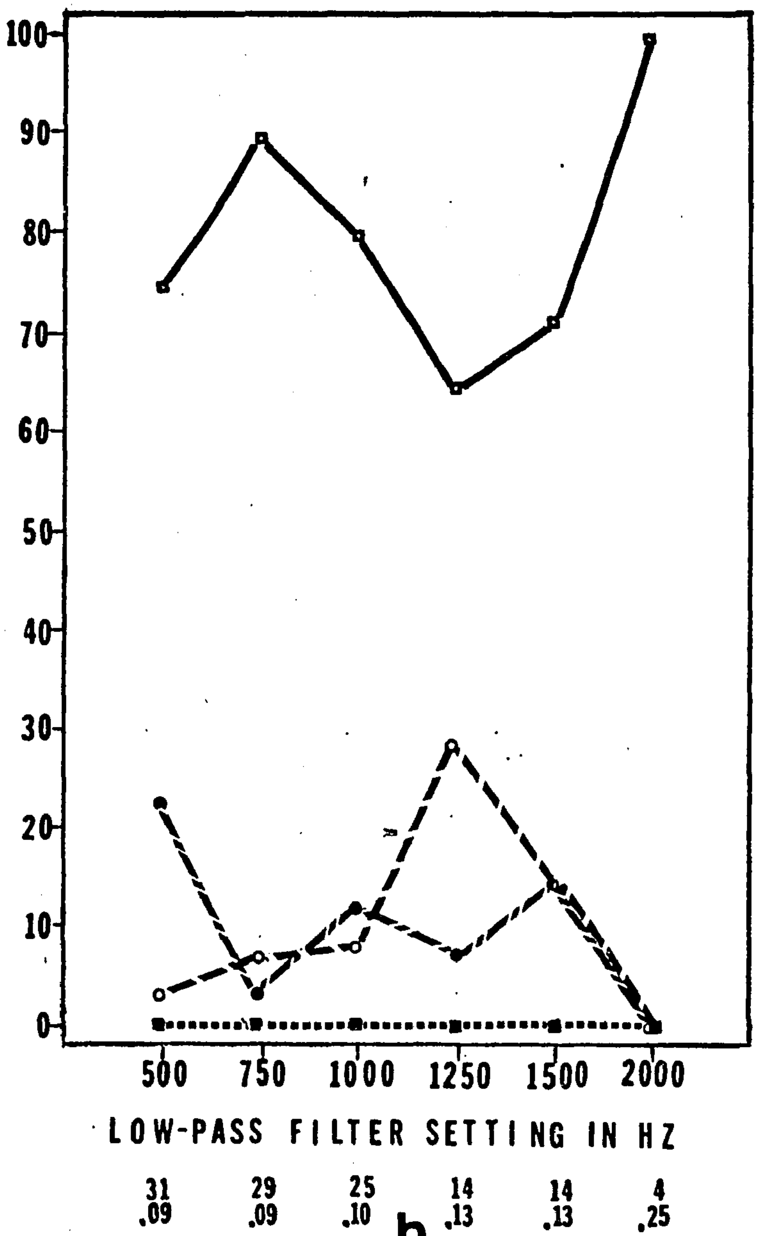
There appear to be more nasality confusions for terminal nasals than for terminal plosives. The nasality contrast for nasals was always the homorganic voiced plosive, while the nasality contrast for plosives was presented in the homorganic nasal. For voiceless plosives, this involves two features, and this may account for the stronger tendency to hear a nasal as a plosive than to hear a plosive as a nasal. The proportion of place confusions is always rather high for the nasals. For all test conditions, listeners confuse nasals with each other more often than with other classes.

Figure 13 illustrates the feature confusions for initial affricates. Since affricates have stop onsets and fricative releases, there was some question as to their

Figure 12. Analysis of feature confusions for initial (12a) and terminal (12b) nasal consonants. Number of errors is shown below each filter setting. The lower row of numbers is an upper bound on the expected standard error for the measured errors at that filter setting. The upper bound corresponds to the standard error for p.=0.5.

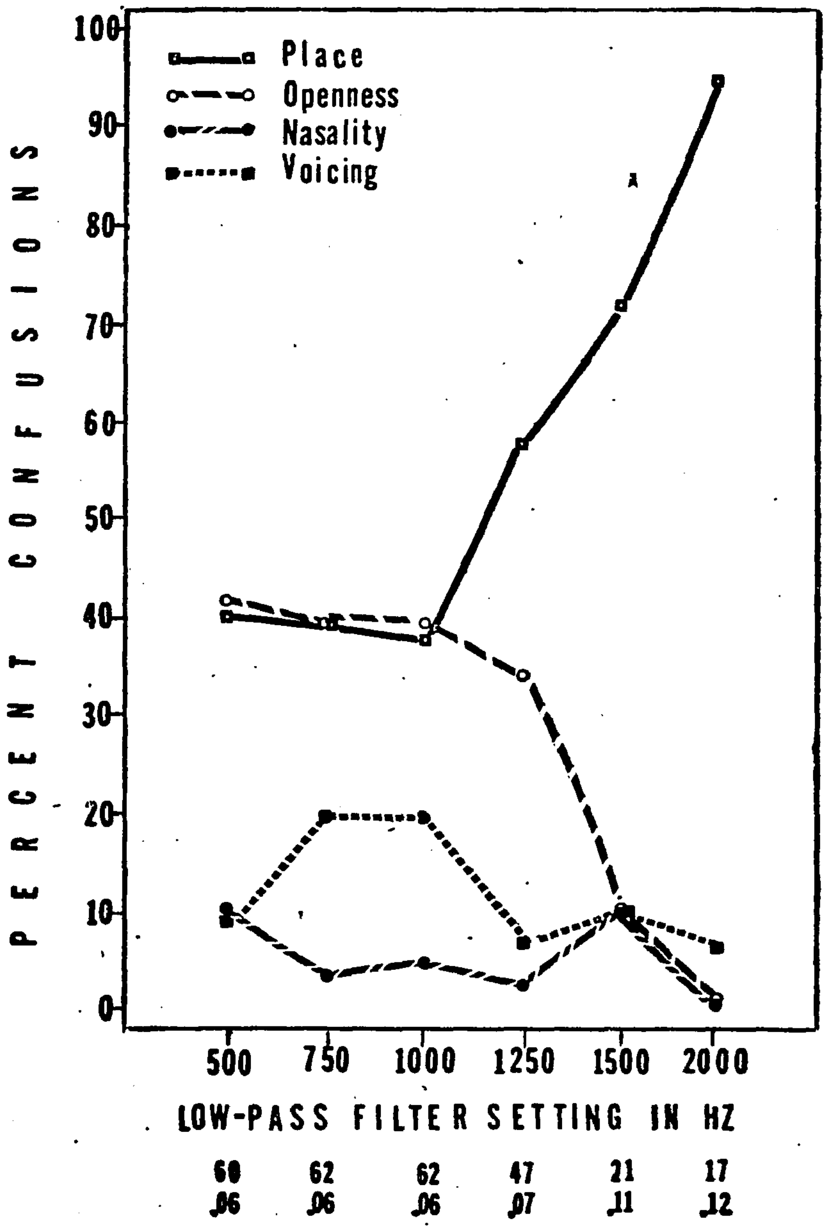

Place
Openness
Nasality
Voicing
PERCENT CONFUSIONS
100
90
80
70
60
50
40
30
20
10
0
500
750
1000
1250
1500
2000
LOW-PASS FILTER SETTING IN HZ
27
.10
23
.10
24
.10
19
.11
10
.16
6
.20

a


100
90
80
70
60
50
40
30
20
10
0
500
750
1000
1250
1500
2000
LOW-PASS FILTER SETTING IN HZ
31
.09
29
.09
25
.10
14
.13
14
.13
4
.25

b

Figure 13. Analysis of feature confusions for initial affricates. Number of errors is shown below each filter setting. The lower row of numbers is an upper bound on the expected standard error for the measured errors at that filter setting. The upper bound corresponds to the standard error for p.=0.5.

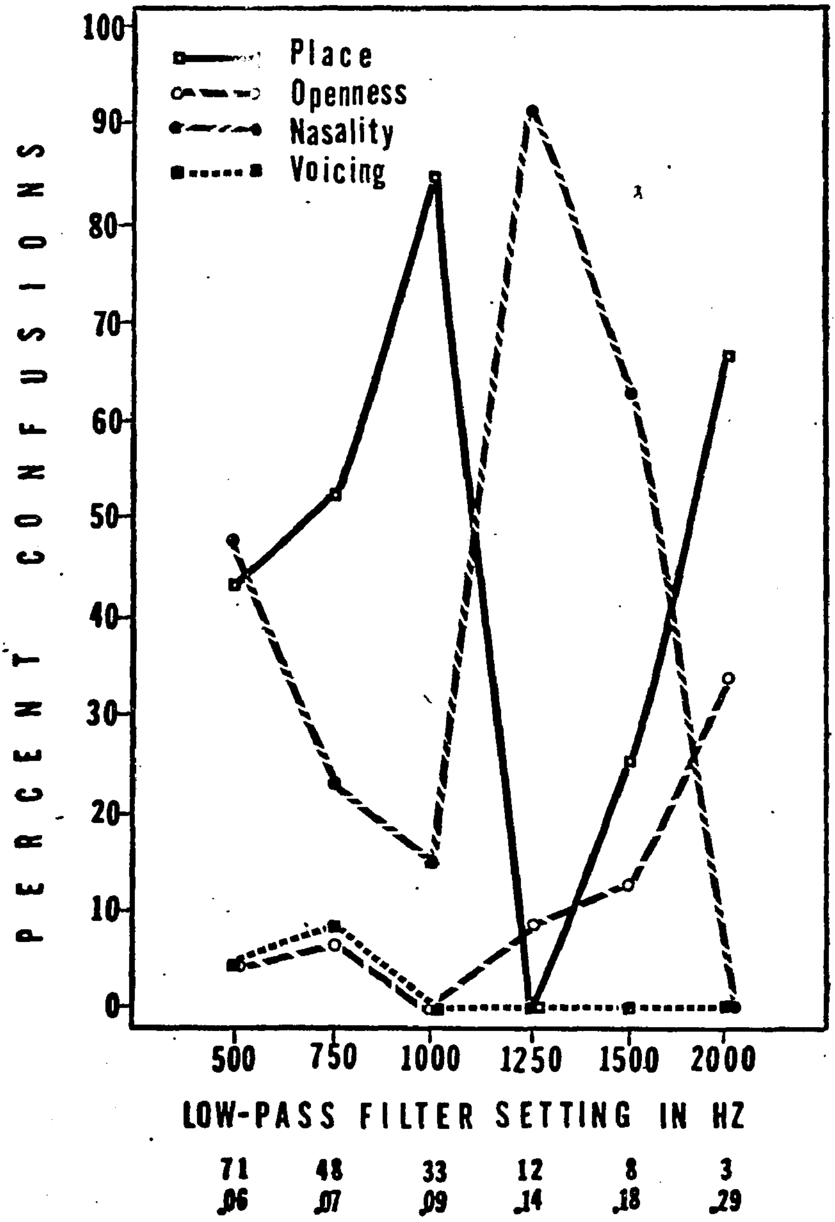

Place
Openness
Nasality
Voicing
PERCENT CONFUSIONS
100
90
80
70
60
50
40
30
20
10
0
500
750
1000
1250
1500
2000
LOW-PASS FILTER SETTING IN HZ
60 62 62 47 21 17
.06 .06 .06 .07 .11 .12

proper place in the phonemic model developed to establish the minimal contrasts. It was decided to classify them as stops. Therefore, place contrasts for the affricates were plosives, and openness contrasts were fricatives. If the affricates had been classified as fricatives, the predominant plosive-affricate substitutions would have been labeled openness confusions. But the general tendency for consonants is a predominance of place errors, and one would therefore expect a predominance of place errors on affricates. The decision to classify them as stops resulted in a majority of "place" confusions; and therefore, it appears that the choice made was the better one.

Above 1000 Hz there is a large majority of place errors; that is, affricates were confused with plosives most of the time, and were rarely confused with fricatives. However, at the 500, 750 and 1000 Hz settings, there was an equal proportion of confusions with fricatives and plosives. This suggests that the "plosive" quality of the affricates is somehow lost when frequencies above 1000 Hz are filtered out. It could be that the sudden onset characteristic of the affricates was too weak or too high in frequency to be transmitted under these low-pass conditions. In addition, there appears to be a greater proportion of voicing errors for these low-pass conditions. In this respect, the affricates resemble fricatives more than plosives.

Figure 14 plots the confusions of the initial semivowels. (No terminal semivowels were tested.) The data for the semivowels is the most erratic of all the phoneme classes.

Figure 14. Analysis of feature confusions for initial semivowels. Number of errors is shown below each filter setting. The lower row of numbers is an upper bound on the expected standard error for the measured errors at that filter setting. The upper bound corresponds to the standard error for p.=0.5.


Place
Openness
Nasality
Voicing
PERCENT CONFUSIONS
100
90
80
70
60
50
40
30
20
10
0
500
750
1000
1250
1500
2000
LOW-PASS FILTER SETTING IN HZ
71
.06
48
.07
33
.09
12
.14
8
.18
3
.29

There is a uniformly low proportion of voicing errors, which is not surprising since there are no voiceless semivowels with which they could be confused. (Voicing contrasts were voiceless fricatives--a two-feature contrast.) Again we find support for the theory that features are confused one at a time. Any semivowel-voicing confusion would involve both voicing and openness.

What is most remarkable about the semivowel data is the sharp reversal of the preponderance of place and nasality errors for the 1250 and 1500 Hz conditions. When all frequencies above 1250 Hz are removed, the semivowels are heard as nasals in more than 90% of the error responses. This perceptual switch to nasals has been observed in studies of synthetic semivowel stimuli. Lisker (1957) found that when the first formants for the synthetic semivowels were set too low, they were sometimes heard as nasals. A similar observation was reported by O'Connor _et_ _al_. (1957). In the present data, the switch to nasal responses does appear when the frequencies above 500 Hz are removed. This could cause a perceptual lowering of the first formant. However, the nasal-for-semivowel responses with the 1250 and 1500 Hz settings cannot be explained in this way. While exact spectral locations of the anti-resonances which characterize the nasals are not known for the speaker in this experiment, it is possible that the removal of that portion of the spectrum around 1250 Hz gives the impression of such an anti-resonance and causes semivowels to be heard as nasals.

The relative intelligibility of each of the phoneme classes is shown in Figures 15 and 16. In the initial position, semivowels are the most intelligible. Semivowels are among the most vowel-like of consonants, produced with a relatively open vocal tract and having a harmonic rather than a noise structure. Therefore, it is understandable that they would share the vowels' relatively high intelligibility, and, that like the vowels, they should lose their superiority at the 500 Hz filter setting.

After semivowels, nasals are most intelligible. These too are somewhat vowel-like in their harmonic structure.

Plosives rank third in intelligibility, being more intelligible for the lowest frequency conditions in the terminal position. As shown in Figure 10, there was a greater proportion of place errors for the terminals in these conditions, and it appears that more place errors are associated with greater intelligibility.

In the initial position, fricatives are less intelligible than plosives under all but the two poorest listening conditions. Fricatives are far less intelligible in the terminal position than in the initial. Their overall scores for the terminal position may help to explain the preponderance of openness errors seen in Figure 11. That is, more openness errors are associated with lower intelligibility, while more place errors are associated with higher intelligibility.

Affricates are clearly the least intelligible of the

Figure 15. Relative intelligibility of five initial consonant phoneme classes.

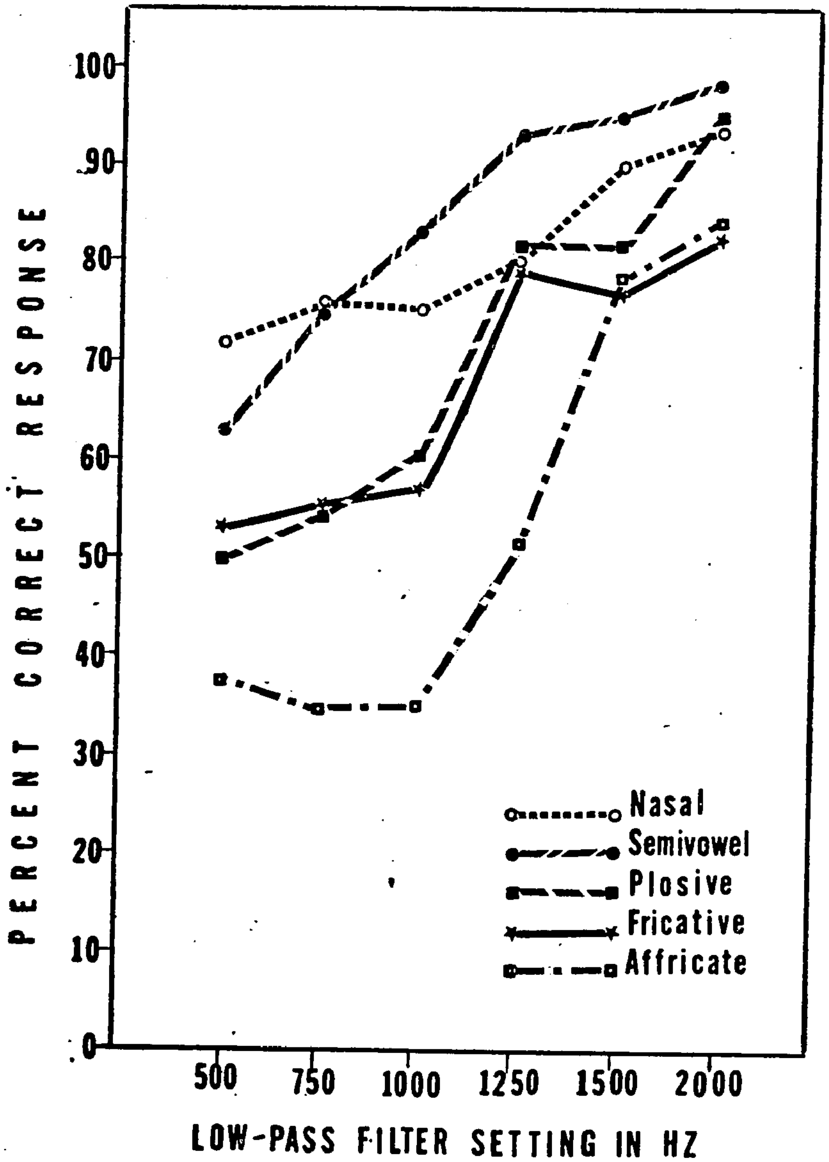
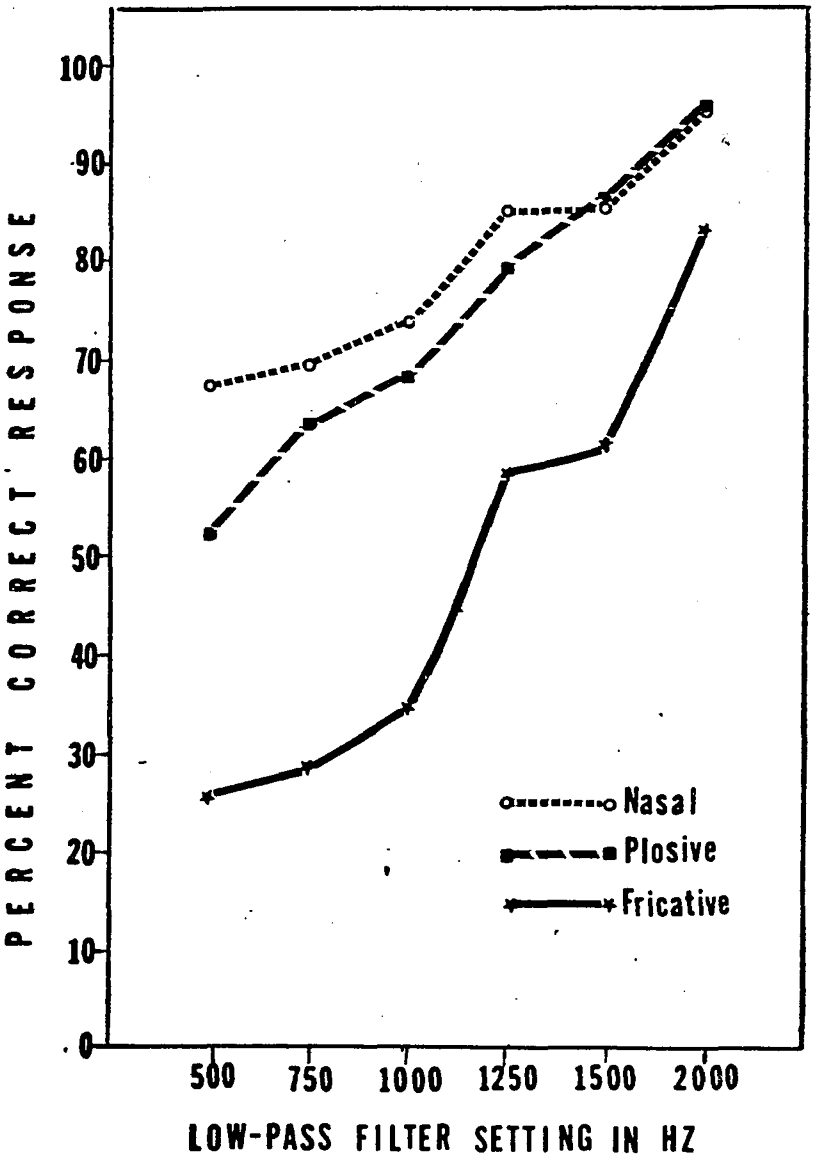

PERCENT CORRECT RESPONSE
100
90
80
70
60
50
40
30
20
10
0
500
750
1000
1250
1500
2000
LOW-PASS FILTER SETTING IN HZ
Nasal
Semivowel
Plosive
Fricative
Affricate

Figure 16. Relative intelligibility of three terminal consonant phoneme classes.


PERCENT CORRECT RESPONSE
100
90
80
70
60
50
40
30
20
10
0
500
750
1000
1250
1500
2000
LOW-PASS FILTER SETTING IN HZ
Nasal
Plosive
Fricative

consonants. It is interesting to note that affricates also exhibited a relatively low proportion of place confusions.

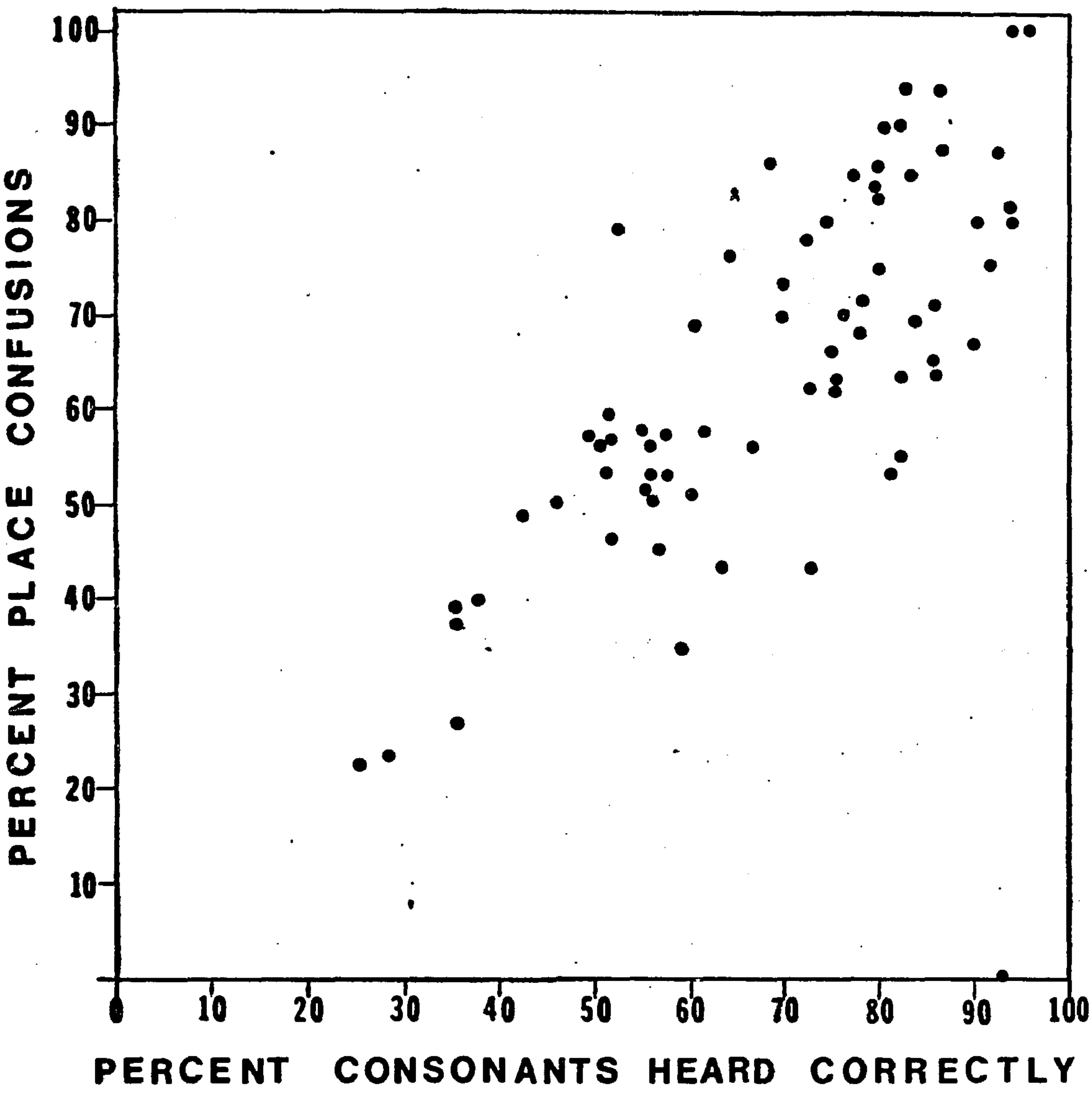
In the terminal position, only nasals, plosives, and fricatives were tested. Their rank order of intelligibility is the same as that found for the initial position.

The feature confusion analyses presented above seems to indicate that whenever a class of phonemes is relatively intelligible, the proportion of place errors is relatively high. It appears that this trend would become evident if the proportion of place errors for each class of phoneme were plotted against the percentage of that class of phonemes heard correctly, since the intelligibility of each consonant class is somewhat different for each test condition.

Figure 17 shows the relationship between intelligibility and proportions of place confusions for all classes of consonants. These classes are the twelve plotted in Figures 8 through 14 (voiced, voiceless, plosive, nasal, etc.). For each class of consonant, the data for conditions with scores less than 95.0% were plotted. Data points for classes and conditions with less than 5.0% error appear to be random and would probably be of no interest in any future applications of the PD Test and procedure.

Inspection of the scatter plot for consonants shows a linear relationship with a positive correlation between the proportion of place confusions and the intelligibility score. For linear data the Pearson-r expresses the

Figure 17. Proportion of place confusions for each class of consonants plotted against percent correct for that class.


100
90
80
70
60
50
40
30
20
10
0
10
20
30
40
50
60
70
80
90
100
PERCENT PLACE CONFUSIONS
PERCENT CONSONANTS HEARD CORRECTLY

relationship between the two parameters. For the consonant data plotted in Figure 17, the Pearson-r is .65, a substantial correlation.

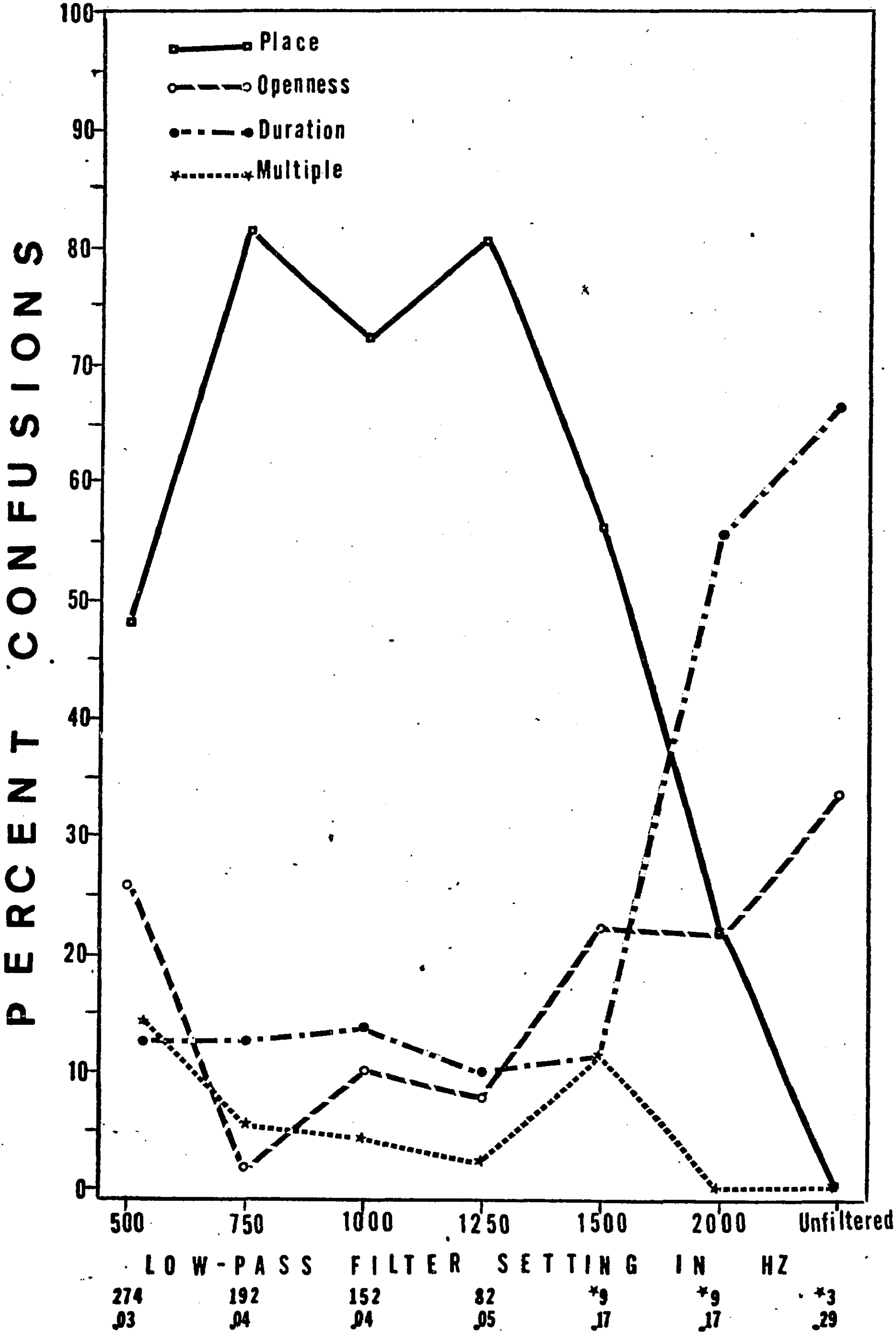
The Vowels.- In the analysis of the vowel confusions, three features were considered--place, openness, and duration. The fourth, multiple-feature contrast, was needed to provide a probability of error equal to that in the consonant tests.

Figure 18 presents the proportion of each of the four contrasts for vowel errors. Place errors are clearly the most frequent. Duration confusions are second in frequency for most conditions. Multiple-feature confusions are the least frequent for all conditions but 500 and 750 Hz.

The scatter and reversals for both the higher frequency conditions and the unfiltered condition may result because such a small number of errors are analyzed. Above 1250 Hz the errors for each condition were less than 2.0%. For the more detailed analyses which follow, any condition with less than 2.0% error was not plotted. While the data at these settings appears erratic, it is interesting to see that there were no multiple-feature confusions at the 2000 Hz or the unfiltered condition. For these conditions, most errors involved duration. It may be that there is sufficient formant frequency information under these conditions to prevent most place and openness errors, and, hence, the few errors which inevitably occur with a sample of this size primarily involve duration. Actually, seven of the nine errors for the 2000 Hz setting occurred in response to

Figure 18. Proportion of vowel errors involving each of four contrasts for each filter condition. Total number of vowel stimuli presented was 3360. Number of vowel errors for each condition is shown below filter setting. The lower row of numbers is an upper bound on the expected standard error for the measured errors at that filter setting. The upper bound corresponds to the standard error for p.=0.5

* Less than 2% error.


Place
Openness
Duration
Multiple
PERCENT CONFUSIONS
100
90
80
70
60
50
40
30
20
10
0
500
750
1000
1250
1500
2000
Unfiltered
LOW-PASS FILTER SETTING IN HZ
274
.03
192
.04
152
.04
82
.05
*9
.17
*9
.17
*3
.29

the /i/ stimulus. Two were place confusions--/u/ for /i/, and four were duration confusions--/ɪ/ for /i/. Confusions of /i/ occur because its second formant is usually at about 2300 Hz, that is, above the filter cut-off. In addition, it appears that under relatively good listening conditions, vowels are usually confused with those nearby on the vowel chart. This would explain the /ɪ/-/i/ substitution: these two vowels do not differ very much from each other in either of the first two formant frequencies.

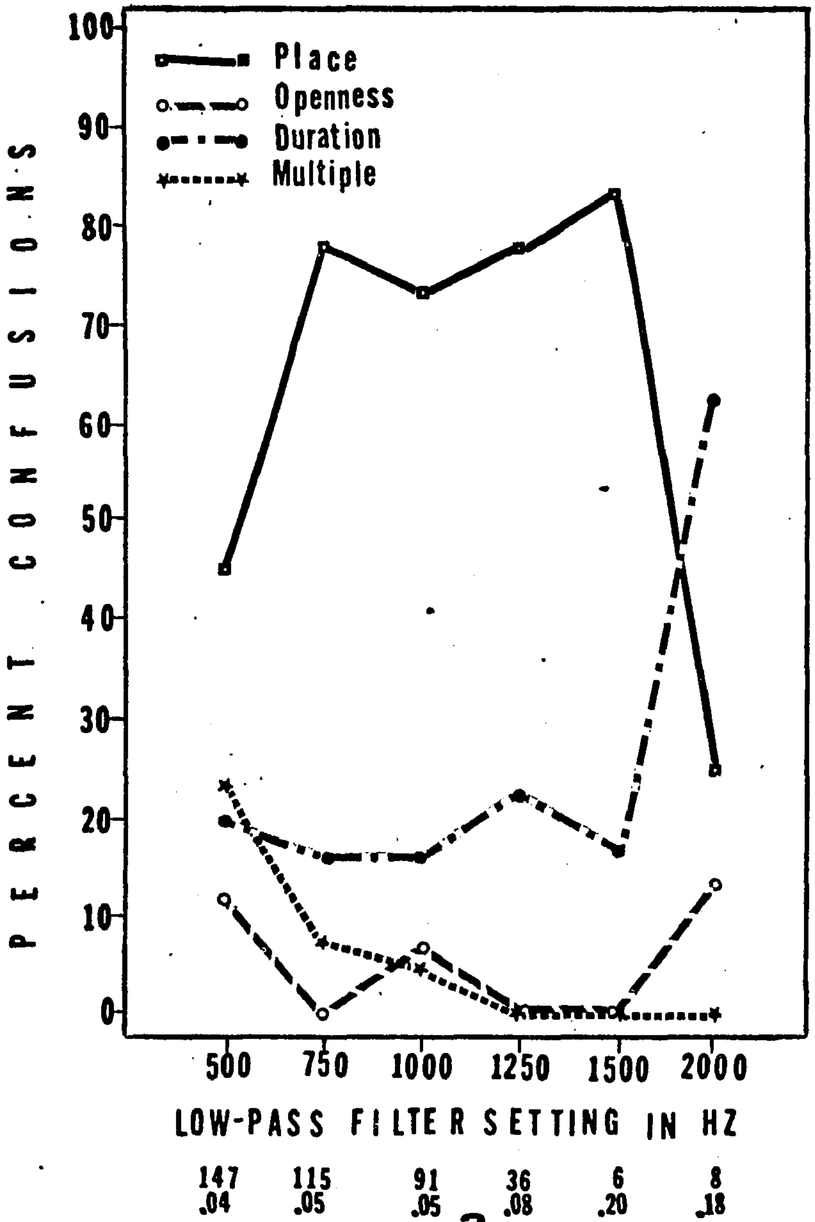
Figure 19 gives the analysis according to vowel duration. The reversals at 2000 Hz for the long vowels may be because of the low error rate (3.0%). The major difference between the long and short vowel responses appears to be the higher rate of duration confusions for long vowels. That is, the filtering tends to cause long vowels to be heard as short vowels more often than short vowels as long ones.

The high proportion of openness confusions for the short vowels at the 500 Hz setting is primarily a matter of the frequent responses of /ɪ/ to the /ɛ/ stimulus, and /ɪ/ or /ʊ/ to the /ʌ/ stimulus. It seems that elimination of the upper portion of the first formant bands of /ɛ/ and /ʌ/ caused these vowels to be heard as vowels with lower first formant frequencies.

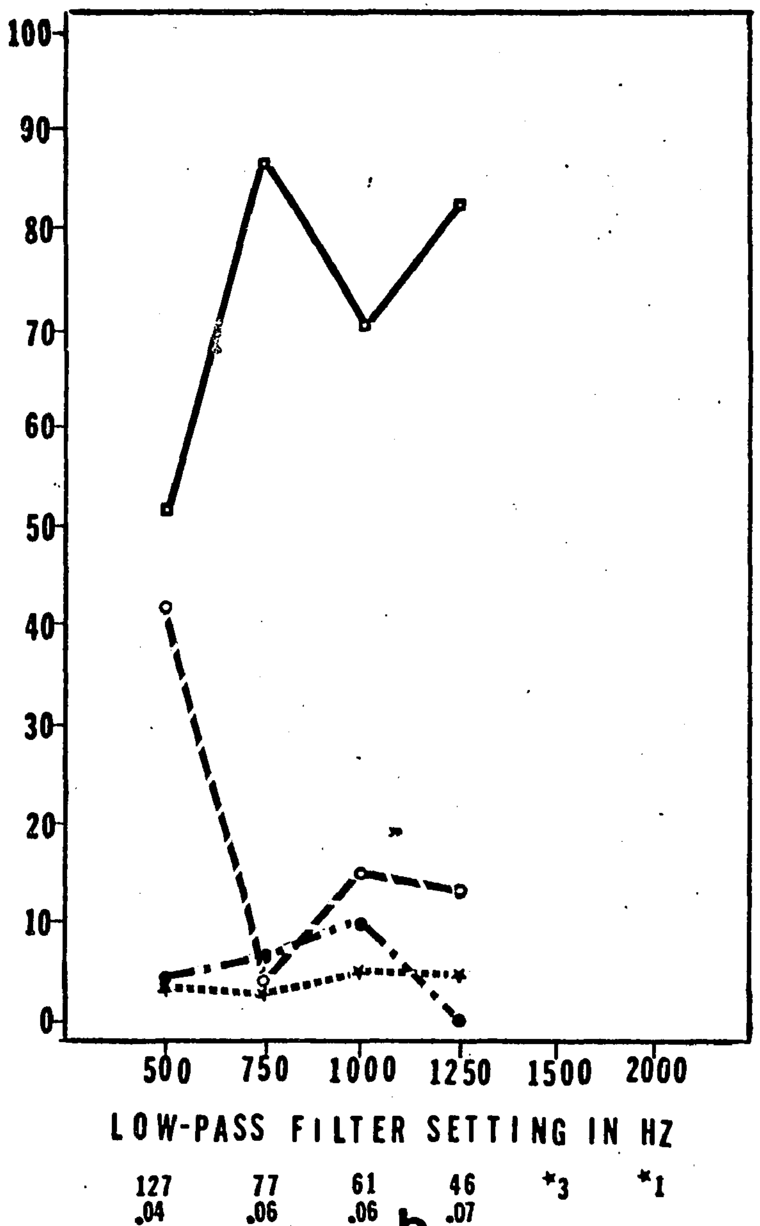
The analyses of confusions for high, medium, and low vowels are shown in Figures 20a, 20b, and 20c, respectively. For most of the conditions, vowels with medium openness exhibit more openness confusions, and correspondingly fewer

Figure 19. Analysis of feature confusions for long (19a) and short (19b) vowels. Number of errors is shown below each filter setting. Lower row of numbers is an upper bound on the expected standard error for the measured errors at that filter setting. The upper bound corresponds to the standard error for p.=0.5.

* Less than 2% error.


PERCENT CONFUSIONS
100
90
80
70
60
50
40
30
20
10
0
Place
Openness
Duration
Multiple
500
750
1000
1250
1500
2000
LOW-PASS FILTER SETTING IN HZ
147
.04
115
.05
91
.05
36
.08
6
.20
8
.18

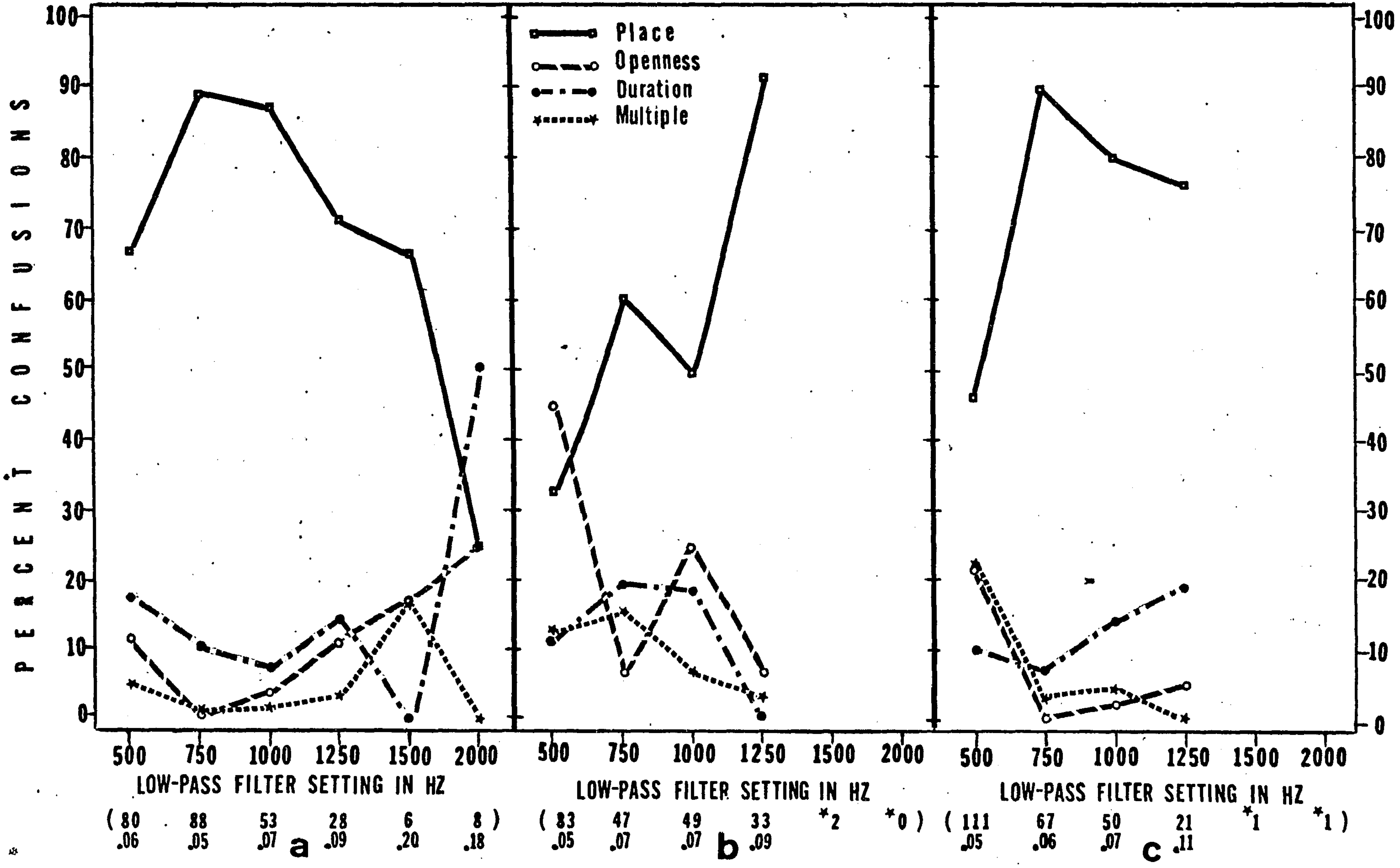
a


100
90
80
70
60
50
40
30
20
10
0
500
750
1000
1250
1500
2000
LOW-PASS FILTER SETTING IN HZ
127
.04
77
.06
61
.06
46
.07
*3
*1

b

Figure 20. Analysis of feature confusions for high (20a), medium (20b), and low (20c) vowels. Number of errors is shown below each filter setting. The lower row of numbers is an upper bound on the expected standard error for the measured errors at that filter setting. The upper bound corresponds to the standard error for p.=0.5.

* Less than 2% error.


PERCENT CONFUSIONS
0
10
20
30
40
50
60
70
80
90
100
Place
Openness
Duration
Multiple
500
750
1000
1250
1500
2000
LOW-PASS FILTER SETTING IN HZ
(80 88 53 28 6 8)
.06 .05 .07 .09 .20 .18
a
(83 47 49 33 *2 *0)
.05 .07 .07 .09
b
(111 67 50 21 *1 *1)
.05 .06 .07 .11
c

place confusions, than do the high or low vowels. Thus, the increase of openness confusions for medium vowels is actually a reflection of the same data seen for the short vowels discussed above, the vowel /ɛ/ being both short and of medium openness.

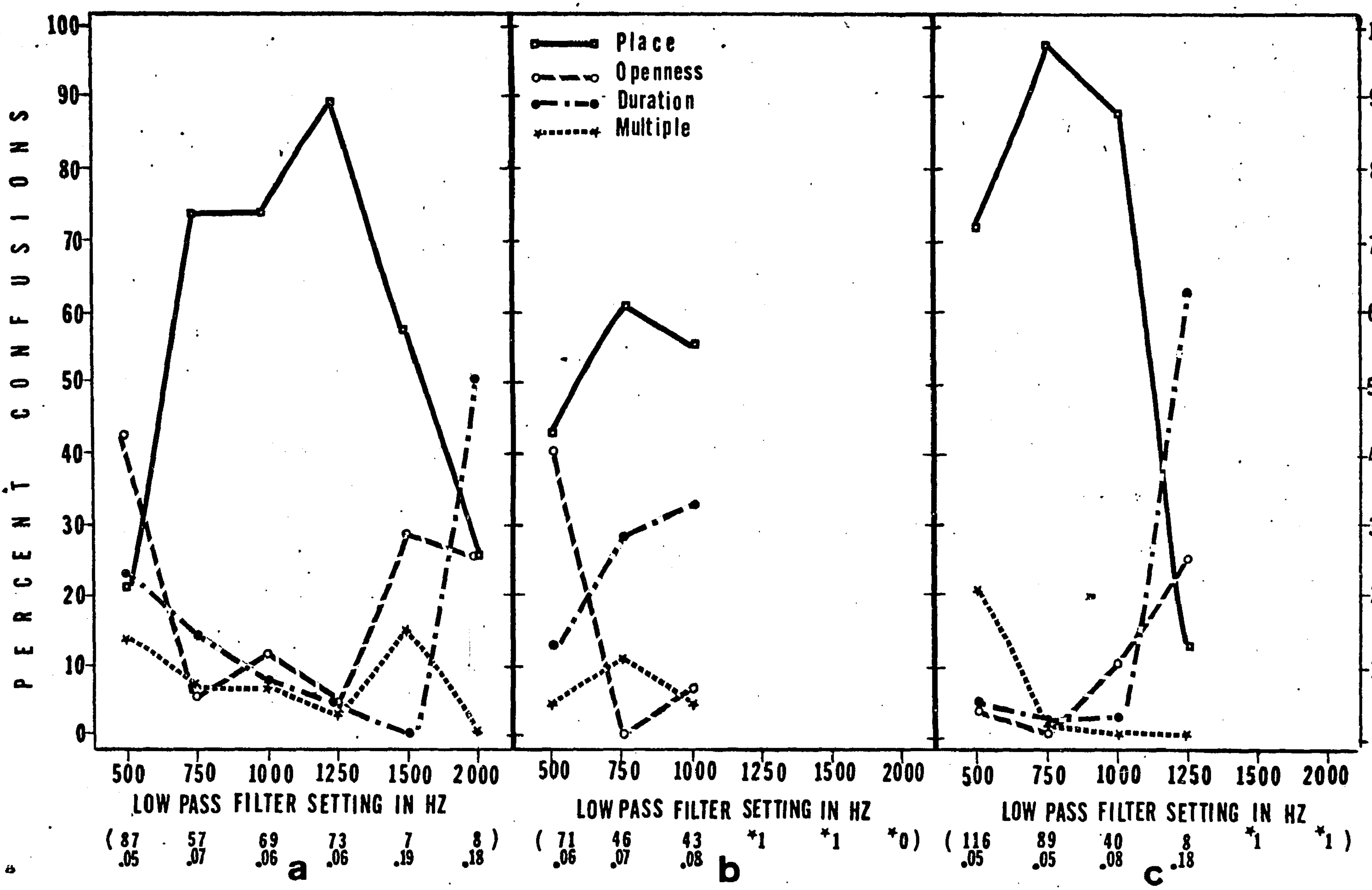
Figures 21a, 21b, and 21c show the proportions of confusions for front, central, and back vowels. There are very few errors above the 1000 Hz setting for both central and back vowels. For front vowels there are still many errors at the 1250 Hz setting. This is because the second formants for front vowels are all above 1250 Hz. The second formant frequency is the cue to place, and there is a very high proportion of place errors for front vowels at the 1250 Hz setting. The peak of place errors for central and back vowels occurs at the 750 Hz setting, since most of the second formants are above this frequency for these vowels.

For all vowel classes, as scores become very high, multiple errors decrease to zero, while duration errors increase. For the 2000 Hz setting and the unfiltered condition, the majority of substitutions were /ɪ/ for /i/ and /ʌ/ for /ɑ/. In both cases, the substituted sound was that nearest the stimulus on the vowel chart.

Figure 22 plots the articulation gain functions for front, central, and back vowels. When any frequencies between 1250 and 2000 Hz are removed, front vowels are least intelligible. At these frequency settings, central and back vowels are near their maximum intelligibility, with a very

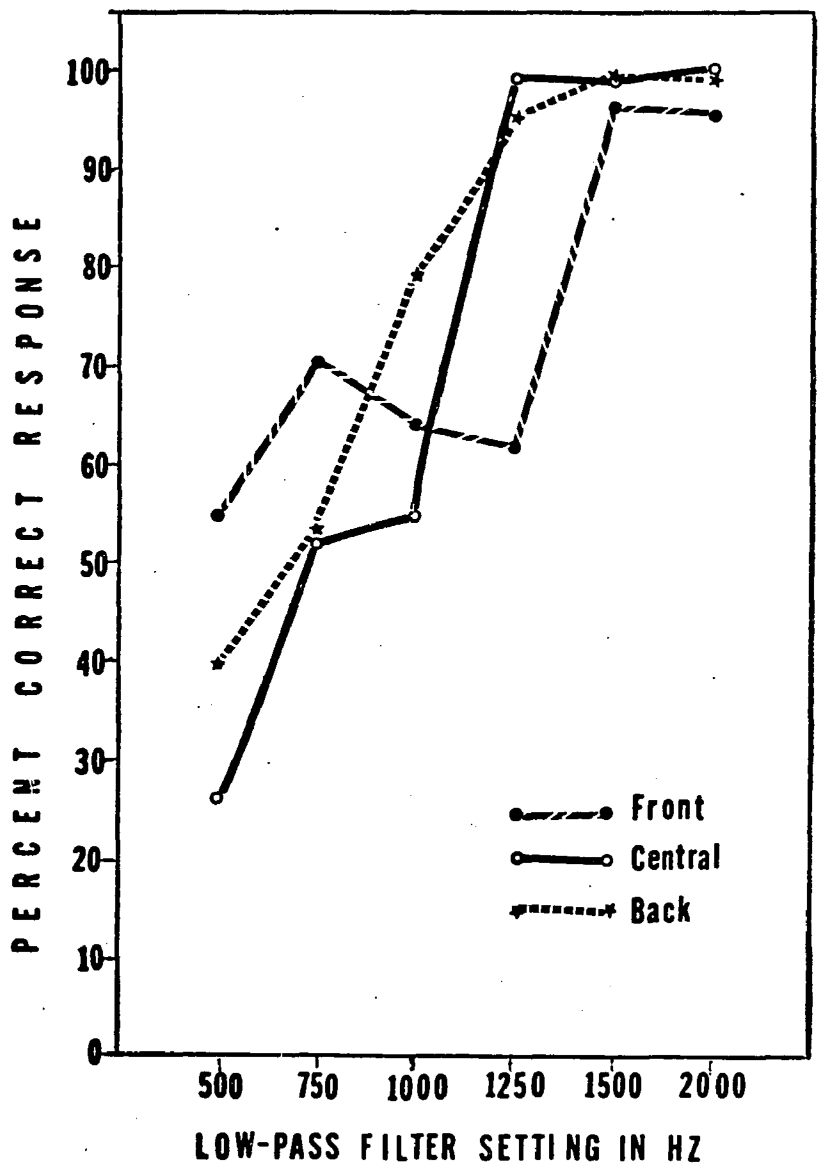
Figure 21. Analysis of feature confusions for front (21a), central (21b), and back (21c) vowels. Number of errors is shown below each filter setting. The lower row of numbers is an upper bound on the expected standard error for the measured errors at that filter setting. The upper bound corresponds to the standard error for p.=0.5.

* Less than 2% error.



PERCENT CONFUSIONS
100
90
80
70
60
50
40
30
20
10
0
Place
Openness
Duration
Multiple
500
750
1000
1250
1500
2000
LOW PASS FILTER SETTING IN HZ
(87 .05 57 .07 69 .06 73 .06 7 .19 8 .18)
a
500
750
1000
1250
1500
2000
LOW PASS FILTER SETTING IN HZ
(71 .06 46 .07 43 .08 *1 *1 *0)
b
500
750
1000
1250
1500
2000
LOW PASS FILTER SETTING IN HZ
(116 .05 89 .05 40 .08 8 .18 *1 *1)
c

Figure 22. Relative intelligibility of front, central, and back vowels.


PERCENT CORRECT RESPONSE
100
90
80
70
60
50
40
30
20
10
0
500
750
1000
1250
1500
2000
LOW-PASS FILTER SETTING IN HZ
Front
Central
Back

slight amount of information to be gained above 1250 Hz for back vowels. The 1000 Hz setting appears to be the cross over point for relative intelligibility of these three classes of vowels. With this setting, the central vowels are the least intelligible of all, but gain over 50% of their maximum intelligibility from information between 1000 and 1250 Hz.

Below the 1000 Hz setting, front vowels are the most intelligible, followed by the back vowels, and then central vowels. The curve for the front vowels exhibits an unusual peak at 750 Hz, which is 6.2% better than the score at 1000 Hz, and 8.3% better than the score at 1250 Hz. Apparently there is no additional information between 750 and 1250 Hz which can contribute to the correct perception of front vowels. Instead it appears that the addition of these frequencies is detrimental to the perception of the front vowels, while it is vital to the perception of central and back vowels. This is easily explained when we consider the formants of front vowels. All of the first formants are below 750 Hz in frequency, while all of the second formants are above 1250 Hz. The inclusion of this portion of the spectrum between the formants causes confusions with those vowels with first formants similar to the stimulus, and second formants between 750 and 1250 Hz. Thus, front vowels are frequently confused with back vowels having similar degrees of openness. The most frequent confusions for these conditions were: /u/ for /i/, /ɑ/ for /æ/, and /ʌ/ for /ɛ/, all place errors.

Figure 22 shows that almost all of the information necessary for correct perception of the central vowels is found in the two frequency bands, 500-750 Hz and 1000-1250 Hz. These bands correspond to the first two formant regions for the two central vowels tested, /ʌ/ and /ɝ/. The back vowels seem to gain intelligibility steadily from the addition of any frequencies from 500 to 1250 Hz since this portion of the spectrum contains all of the second formants for the back vowels.

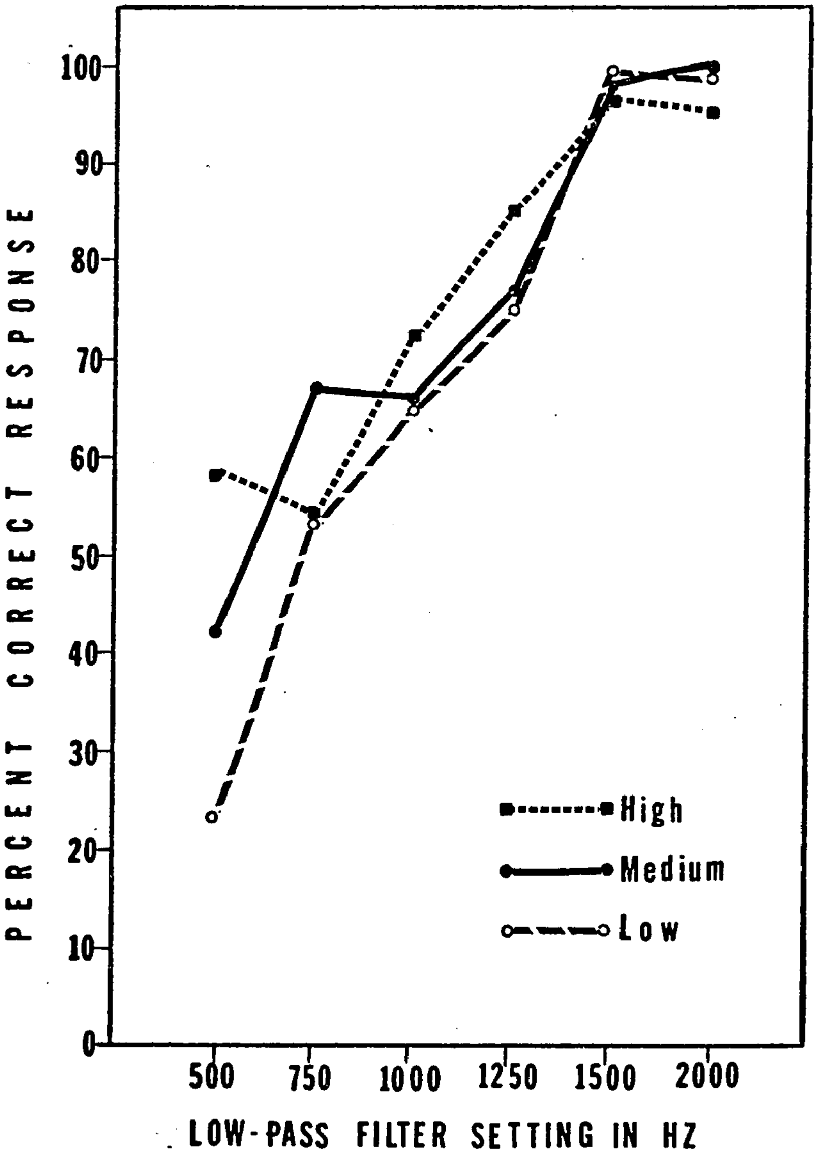
Figure 23 plots the relative intelligibility of high, medium, and low vowels. The lower scores for the high vowels at 1500 Hz and 2000 Hz are a reflection of the data for the front vowels discussed above, that is, the errors in perception of the vowel /i/, which is high and front. The leveling off short of maximum intelligibility for the long vowels is also seen in Figure 24, and again it represents errors for /i/.

The high vowels improve steadily from 750 to 1500 Hz. The relatively greater intelligibility of the high vowels at the 500 Hz setting is primarily a reflection of the vowels /i/, /ɪ/, and /u/. Since two of these are front vowels, the same superiority was shown for the front vowels in Figure 22.

The low vowels show a steady increase in score as frequencies from 500 to 1500 Hz are added, probably because this portion of the spectrum includes all of the second formant information for the low vowels.

It is also suggested by Figure 23 that the spectrum

Figure 23. Relative intelligibility of high, medium, and low vowels.

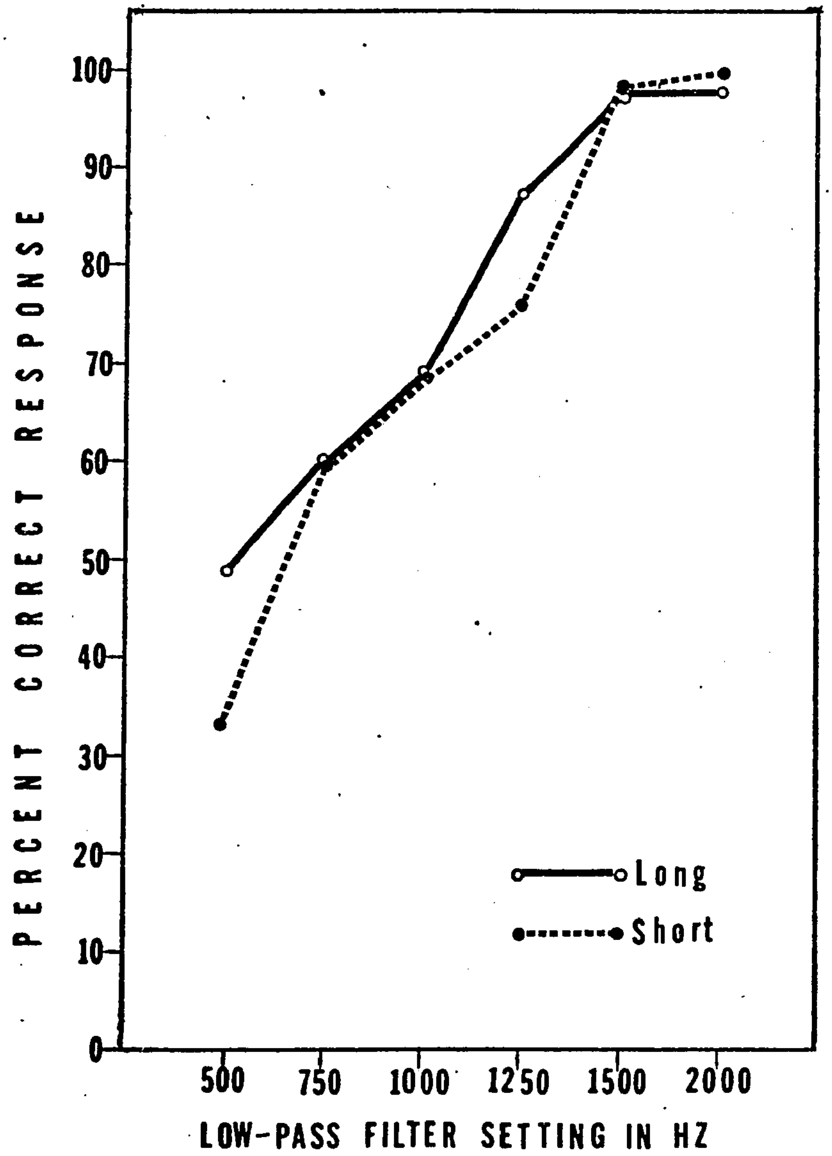

PERCENT CORRECT RESPONSE
0
10
20
30
40
50
60
70
80
90
100
500
750
1000
1250
1500
2000
LOW-PASS FILTER SETTING IN HZ
High
Medium
Low

between 750 and 1000 Hz does not add to the intelligibility of the medium vowels, and may, in fact, be detrimental. The reversal is actually found only for the vowels /ɛ/ and /ɝ/. For both vowels the first formants are well below 750 Hz, and the second fromants are well above 1000 Hz. For /ɛ/, the addition of the frequencies between 750 and 1000 Hz caused a slight increase of openness confusions, For /ɝ/, the addition of these frequencies caused an increase of both openness and place confusions with vowels having formants in or near this portion of the spectrum.

Figure 24 compares the articulation gains for short and long vowels. Three of the six points plotted are overlapping or close. The difference at 500 Hz has already been established as a matter of relatively high intelligibility for /i/ and /u/, which are both high and long. The lower intelligibility of the long vowels at 2000 Hz is a reflection of the less than perfect intelligibility for /i/, also seen in the front and high vowel curves. The superiority of long over short vowels at the 1250 Hz setting appears to be determined primarily by the relatively poor perception of /ɛ/. The 1250 Hz condition produced more than twice the number of place errors for /ɛ/ than did the 1000 Hz setting. This results from the addition of that portion of the spectrum which contains the second formant of /ʌ/, a place contrast to /ɛ/.

In general, the confusions of vowel features can be explained by their formant locations. In addition, when only one vowel is responsible for the predominance of a feature

Figure 24. Relative intelligibility of long and short vowels.

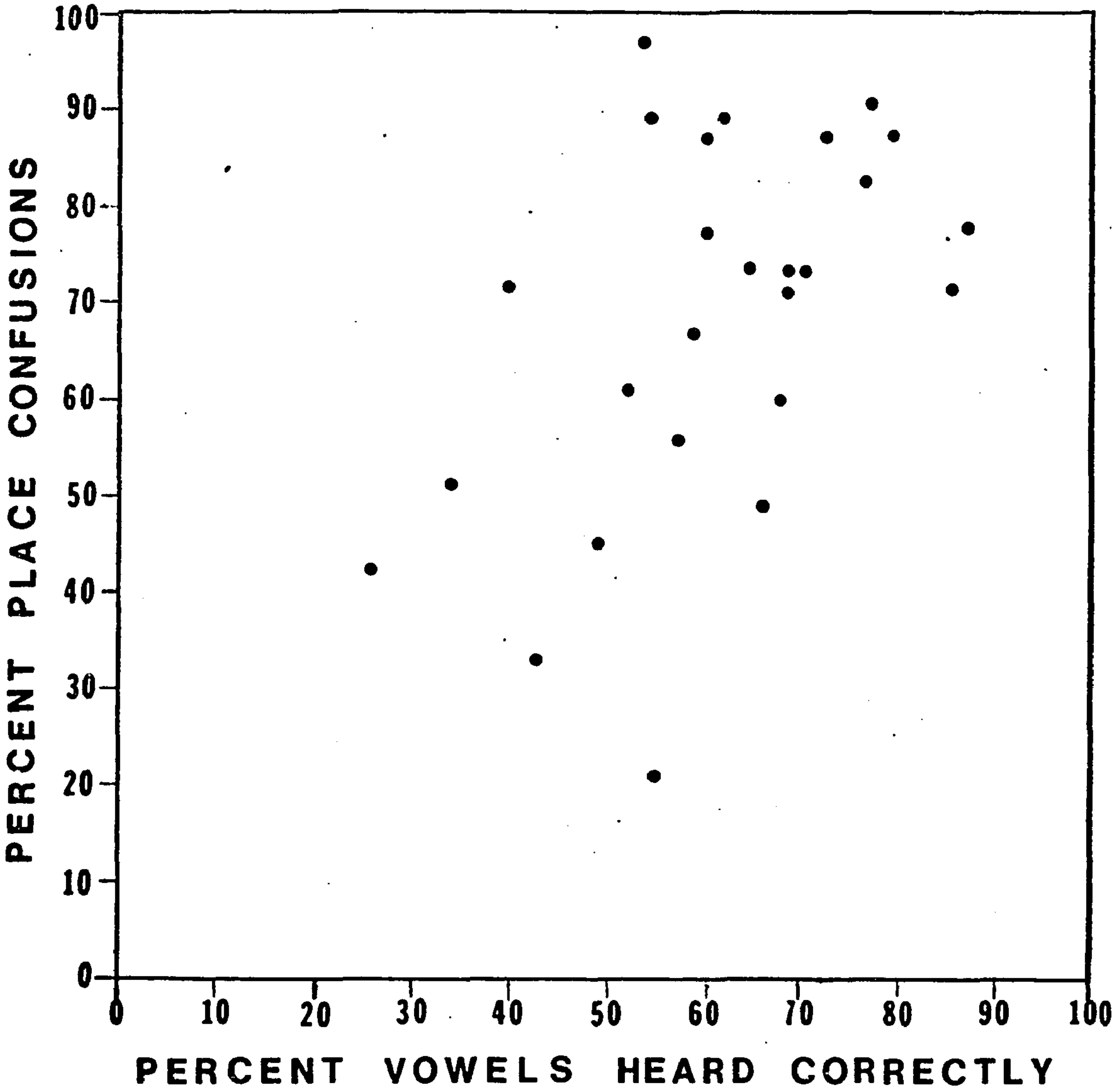

PERCENT CORRECT RESPONSE
0
10
20
30
40
50
60
70
80
90
100
500
750
1000
1250
1500
2000
LOW-PASS FILTER SETTING IN HZ
Long
Short

confusion, the predominance is generally seen in the data for each of this vowel's descriptive classes. For example, if a phenomenon found at a given filter setting for front vowels is similar to the phenomenon found for medium and short vowels, we can assume that the phoneme responsible is /ɛ/.

For all classes of vowels the most evident trend was for a predominance of place confusions. In general, the proportion of place confusions increases with increasing intelligibility but the relationship does not appear to be as strong as that found for consonants. Figure 25 plots the relationship between score and proportion of place confusions for the eight classes of vowels shown in Figures 19, 20, and 21. Entries include only those classes and conditions with scores less than 95.0% There is far more scatter for the vowel plot than there was for the consonant plot. Yet the relationship appears to be linear, and the correlation is positive. The Pearson-r co-efficient is .52, a moderate correlation.

Therefore, any prediction of a predominance of place confusions appears to be less reliable for the vowels than for the consonants. The more random appearance of the vowel data may be attributable to the structure of the vowel contrasts. It was in most cases difficult to find pure one-feature contrasts for the vowels, that is, vowels which differed from the stimuli in the desired feature and not in any other feature. As front vowels become more open they also

Figure 25. Proportion of place confusions for each class of vowels plotted against percent correct for that class.


PERCENT PLACE CONFUSIONS
100
90
80
70
60
50
40
30
20
10
0
0
10
20
30
40
50
60
70
80
90
100
PERCENT VOWELS HEARD CORRECTLY

move back, and as back vowels become more open they generally move forward. This was illustrated in Figure 1. The English vowel chart is more a loop or triangle than it is a rectangle, and thus the vowels are not symmetrically contrasted for place and openness. Differences in openness almost always involve at least some differences in place. For the consonants, on the other hand, clear place differences were usually available. The fact that openness and duration confusions for the vowels also involved at least minor confusions of place of articulation may help to explain the less "predictable" data. It is even possible that some of the confusions which were recorded as openness or duration confusions were in fact place confusions. This seems especially likely for those filter conditions which involved loss of second formant frequencies but preservation of first formant frequencies. For these conditions, listeners should confuse the place feature and not the openness feature.

The Diphthongs.- The percentages of correct perception of the diphthongs for each condition have been given in Table 18 and illustrated in Figure 4.

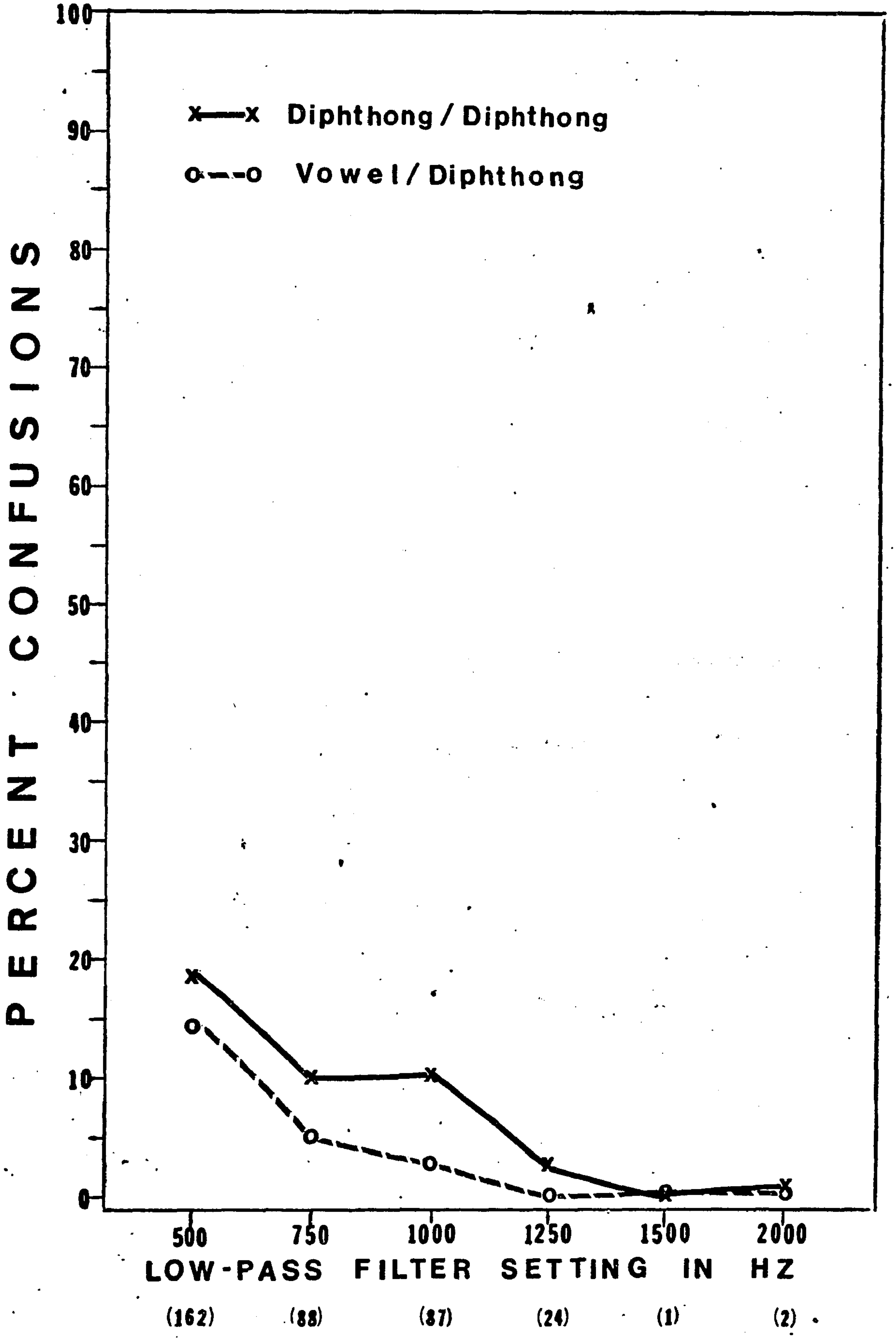
In the development of the test sets, each diphthong was contrasted with as many other diphthongs as possible, and the remaining choices were long vowels, frequently with articulatory positions similar to the onset or offset positions of the diphthong stimulus.

As expected, diphthongs were confused more with other diphthongs than with vowels. Of the 5336 diphthongs presented as foils to the 2080 diphthong stimuli, 334 were chosen as error responses (6.4%). Of the 2024 vowels presented as foils for the diphthong stimuli, 72 were chosen as error responses (3.5%). Figure 26 shows the proportions of stimulus diphthongs which were heard as other diphthongs, and the proportion of stimulus diphthongs which were heard as vowels. For all conditions, vowels were chosen in lower proportions.

Figure 27 shows the relative intelligibility of the five diphthongs tested. All five attain perfect or nearly perfect scores when all frequencies below 1500 Hz are present. The diphthongs /aʊ/ and /ɔɪ/ are not affected by elimination of frequencies above 1250 Hz, but /oʊ/, /aɪ/, and /eɪ/ lose between 10 and 20 per cent of their intelligibility. All five diphthongs gain some information from the frequencies between 1000-1250 Hz, and between 500-750 Hz. It appears that less information is contained in the frequencies between 750-1000 Hz.

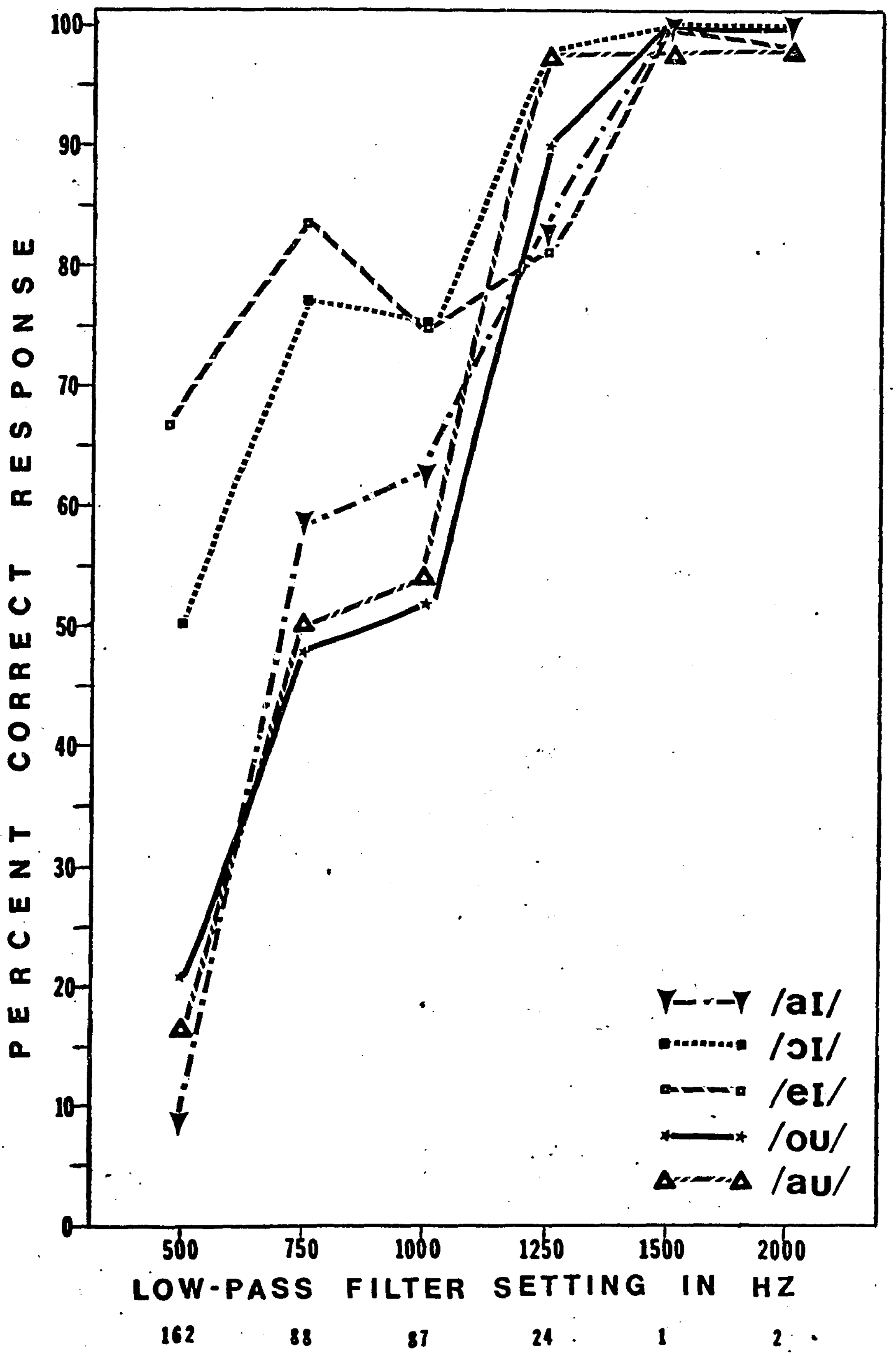
For the two poorest listening conditions, /eɪ/ is the most intelligible, and /ɔɪ/ is second. The other three diphthongs are far less intelligibile under these conditions. When frequencies above 500 Hz are filtered out, the /eɪ/ is heard correctly 66.7% of the time. This is probably because the first formant movement remains intact. For the /ɔɪ/, which is heard correctly half of the time, most of the first formant movement remains. For both /eɪ/ and /ɔɪ/, the entire second formants are above the 500 Hz setting. Thus, they are being perceived according to the single formant vowel movements.

Figure 26. Proportion of diphthongs confused with diphthongs and diphthongs confused with vowels. Total number of diphthongs presented under each filter condition was 240. Number of errors is shown below filter setting.



x——x Diphthong / Diphthong
o- - -o Vowel / Diphthong
PERCENT CONFUSIONS
100
90
80
70
60
50
40
30
20
10
0
500
750
1000
1250
1500
2000
LOW-PASS FILTER SETTING IN HZ
(162)
(88)
(87)
(24)
(1)
(2)

Figure 27. Relative intelligibility of five diphthongs. Total number of diphthong errors for each condition are shown below filter setting.


PERCENT CORRECT RESPONSE
100
90
80
70
60
50
40
30
20
10
0
500
750
1000
1250
1500
2000
LOW-PASS FILTER SETTING IN HZ
162
88
87
24
1
2
/aI/
/ɔI/
/eI/
/ou/
/au/

For the 750 Hz setting, the first formant movement of /ɔɪ/ is completely transmitted, and its score shows a sharp increase.

For the 500 Hz condition, /ɔɪ/ was frequently heard as /oʊ/, which has a very similar first formant movement beginning at around 500 Hz. For the 750 Hz condition, /ɔɪ/ was frequently heard as /aɪ/ or /aʊ/ which have first formant movements beginning at around 750 Hz. The inclusion of frequencies between 750 and 1000 Hz adds nothing to the score for /ɔɪ/. First formant information was at its maximum by 750 Hz, and second formant information is still well above the spectrum transmitted. For both 500 and 750 Hz settings, /eɪ/ was heard as either /ɔɪ/ or /oʊ/, because of the similarity of first formants. For the 500 and 750 Hz conditions, the most frequent errors for /oʊ/ were confusions with /eɪ/. This is a result of the similarity of first formants, as stated above. The second most frequent substitution was /aʊ/, which also has a similar first formant movement, given the spectrum being transmitted.

For the 500 Hz condition, /aʊ/ was confused with all of the other diphthongs. This is because its descending first formant has been removed. For the 750 Hz condition, /aʊ/ was heard as /aɪ/ as frequently as it was heard correctly. This is to be expected because of the identical first formant movements of these two diphthongs. Only the missing second formants could distinguish them. For this same condition, /aɪ/ was frequently heard as /aʊ/. For the settings 1000 and 1250 Hz, /aɪ/ and /aʊ/ were most frequently confused with each other.

The number of vowel contrasts to the diphthongs was limited. However, when a vowel corresponding to the onset or offset position for a stimulus diphthong was given as an alternative, it was clearly the most frequent vowel chosen in error. The vowel /æ/ was frequently chosen for /aɪ/, and /i/ was frequently chosen for /ɔɪ/. These two vowels have formants which resemble the onset and offset formants of the /aɪ/ and /ɔɪ/, respectively.

As shown in Table 18, when all frequencies below 1250 Hz are present, the diphthongs are the most intelligible phonemes. Under such listening conditions, one should expect relatively few diphthong confusions. For poorer listening conditions, that is, when lower frequencies also are removed, those diphthong confusions which do occur are usually those to be expected from their formant spectra. When only very low frequencies are transmitted,for example, below 500 Hz, one can expect the diphthongs to have the poorest overall intelligibility.

The Hard-of-Hearing Listeners

In order to determine the clinical applicability of the Phoneme Differentiation Test and procedure, nine clinic patients were tested. Table 19 gives the age, sex, diagnosis, history, audiometric slope, test ear, and speech reception threshold (SRT) of each clinical subject.

The audiogram obtained for each clinical subject is

TABLE 19

CLINICAL DATA FOR NINE HARD-OF-HEARING SUBJECTS

Subject	Age	Sex	Diagnosis	History	Audiometric Slope[a]	Test Ear	SRT in dB
TO	10	M	bilateral sensori-neural loss	history of hearing loss on both sides of family	15/25/50/50/65/65/45 (G)	L	34
EB	14	M	bilateral sensori-neural loss	premature birth; delayed speech; poor learner	25/25/35/50/60/60/60 (G-F)	R	32
TE	15	M	bilateral sensori-neural loss	loss noticed at age two; etiology unknown	35/25/40/55/70/65/35 (M)	L	28
CO	18	F	bilateral sensori-neural loss	family history of hearing loss	60/60/55/50/50/65/65 (F)	R	42
WB	26	M	no hearing on right; VIII N. damage left	right VIII N. removed; tumor removal and decompression of left N.	20/25/30/50/55/55/70 (G^m)	L	24
JM	40	F	mild bilateral sensori-neural loss	cerebral palsy; visual and motor problems	20/20/50/50/30/20/10 (T)	L	28
SR	40	M	mild bilateral sensori-neural loss	sudden partial loss in right ear after intensive target practice	30/20/20/10/25/65/35 (R-M)	R	10
JS	62	M	bilateral sensori-neural loss	family history of hearing loss; worked in print shop for 30 yrs.	20/20/30/35/40/50/70 (G)	R	18
MH	77	M	bilateral sensori-neural loss	loss for several years; constant over last year; history negative	45/40/30/40/40/60/NR (F^m)	R	32

[a] The seven numbers given as audiometric slope represent the thresholds in dB at 125, 250, 500, 1000, 2000, 4000 and 8000 Hz. NR indicates no response. Letters indicate the general slope of the audiogram.

given in thresholds for seven octave points. The curve is classified on the chart by one of several code letters described by Carhart (1945). A major category is assigned if the audiometric slope falls within ±10 dB of a "curve of best fit" for at least three successive octaves. In the present paper, a modifying superscript indicates a change in the curve. Only part of the descriptive system presented by Carhart is used here. The five major curve types used are described below:

(F) flat: thresholds are approximately equal for all frequencies
(G) gradual downward slope proportionately greater for high frequencies at a slope of 5 or 10 dB per octave
(M) marked downward slope of 15 or 20 dB per octave
(R) rising: rise of curve at 5 or 10 dB per octave
(T) trough: greater loss os middle frequencies of at least 20 dB.

Table 20 presents the full scores on the PB and PD tests, and the scores on each of the three PD test parts. The number of test items for each of the parts is: 88 initial consonants, 52 terminal consonants, and 60 vowels and diphthongs.

A comparison of the PB and PD scores shows that for eight subjects the score on the PD test is the poorer. It would seem, therefore, that the fine discriminations required by the PD test are more difficult than discriminations required by the PB words in open-set testing. This may result from the fact that, as compared with the PD test, there is less difference between a given PB word and others with which it may be confused, or to the fact that the PB test word is

TABLE 20

SCORES ON PB AND PD TESTS FOR NINE HARD-OF-HEARING SUBJECTS

Subject	PB Score[a]	Full PD Score	Diff. PB-PD[b]	Initial Consonants	Terminal Consonants	Vowels	Diphthongs
TO	88% / 96%	73.0%	-15 /-23	66%	64%	95%	85%
EB	92% / 68%	55.5%	-36 /-12	58%	42%	60%	70%
TE	88%	72.5%	-15.5	78%	54%	78%	85%
CO	84%	76.5%	- 8.5	75%	60%	98%	85%
WB	68% /76%	81.5%	+13.5 / +5.5	86%	63%	90%	100%
JM	96%	95.0%	-1	98%	90%	98%	90%
SR	96%	87.0%	-9	82%	83%	100%	95%
JS	92%	79.5%	-12.5	77%	65%	95%	95%
MH	88%	82.0%	-6	80%	71%	95%	95%
Means	88%	78.1%	-9.9[c]	78%	66%	89%	89%

[a] Second entry represents a retest PB score where available.

[b] Differences between PB and PD scores given for both PB scores.

[c] Mean differences calculated by taking average difference for each subject.

more familiar than the words with which it could be confused. The one exception to the trend for higher PB scores was subject WB who suffered acoustic nerve damage. This subject told the examiner that the PD test seemed easier than PB tests because it was more structured, that is, knowing part of the word made discriminations easier.

The chief reason for testing this small sample of hypacusic subjects was to determine the clinical feasibility of the PD test. Certain general statements can be made about this sample and about its relationship to the findings for the larger, normal sample. However, because of the small heterogeneous clinical sample, such comparisons are limited. Any detailed breakdown will necessarily involve a small number of errors, and therefore, there will be a great deal of scatter in the data. The diagnostic value of the PD test can only be established after study of a large, selected sample of hard-of-hearing subjects. However, certain trends manifest by normal subjects when they listened to filtered speech can also be seen for the hard-of-hearing subjects, and these trends will be discussed below.

Table 20 shows that the clinical subjects made an average of 12% more errors for consonants in the terminal word position than in the initial word position. Eight of the nine clinical subjects scored higher on initial consonants. The normal-hearing listeners for the filtered conditions also averaged more errors on the terminal consonants, but the difference between initial and terminal consonant scores was only 5.4% when all conditions were averaged. For the normal group, the greatest difference between initial and terminal consonant scores was 8.6% for the 500 Hz setting.

Table 20 also shows that all of the nine clinical subjects were better at perceiving vowels than consonants. The average difference was 17% in favor of the vowels. The normal listeners also perceived vowels better under all test conditions except the 500 Hz low-pass filter setting. The average difference was 7%. However, for one of the filter conditions (1500 Hz), the difference between consonant and vowel scores was 17%.

In general, it appears that for both clinical and normal-hearing subjects initial consonant scores are better than terminal consonant scores, and vowel scores are better than consonant scores. None of the tested filter conditions caused the normal-hearing group to show the same difference between initial and terminal consonant scores as was shown by the clinical group. However, the 1500 Hz filter condition caused the normal-hearing listeners to show the same proportion of relatively greater intelligibility of vowels over consonants as that found for the clinical group.

There were few errors on diphthongs in the clinical group--between 0 and 3 errors for all but one subject, EB, who made 6 diphthong errors. This subject's score was also the poorest of the group on all other sections of the PD test. The average intelligibility of vowels and diphthongs was the same, 89%.

TABLE 21

FEATURE ANALYSIS OF CONSONANT ERRORS FOR NINE HARD-OF-HEARING SUBJECTS

	Initial Consonants (*n* each subject = 88)					Terminal Consonants (*n* each subject = 52)				
	Percent Correct	Percentages of Error Types				Percent Correct	Percentages of Error Types			
Subject		Place	Open.	Voic.	Nasal.		Place	Open.	Voic.	Nasal.
TO	66.0	57	27	13	3	63.4	58	32	5	5
EB	58.0	54	22	19	5	42.3	60	23	10	7
TE	78.4	21	42	32	5	53.8	67	4	25	4
CO	75.0	27	27	36	9	59.6	48	33	19	0
WB	86.4	50	33	8	8	63.4	53	37	10	0
JM	97.7	50	50	0	0	90.4	60	20	0	20
SR	81.8	44	44	6	6	82.7	44	44	12	0
JS	77.3	50	30	15	5	65.4	44	33	6	17
MH	79.5	28	33	28	11	71.1	20	60	7	7
Means	78.0	43	31	20	6	66.0	53	30	12	5

Table 21 gives the feature analysis of consonant errors in the clinical group. On the average, place errors are most frequent for consonants in both initial and terminal-word positions. The mean proportion of place confusions by the clinical group on initial consonants (43%) is lower than the proportion for the normal-hearing group with any of the filter conditions, but the mean proportion of place confusions for terminal consonants (53%) by the clinical group is similar to that for the normal-hearing group with the 1000 Hz filter setting (51%).

For initial and terminal consonants, openness confusions were second most frequent for both normal and clinical groups. The proportion of initial consonant openness confusions by the clinical subjects (31%) was higher than for any of the filter conditions, but was similar to the proportion of openness confusions for the 1000 Hz setting (29.4%). The average proportion of terminal consonant openness confusions by the clinical group (30%) was similar to that found for the averaged filter conditions (31.2%).

Voicing errors were less frequent than place or openness errors in both groups of subjects, but occurred more frequently in the hard-of-hearing than in the normal group. Although the hard-of-hearing showed proportionally more voicing errors than normals, they did show fewer voicing errors for terminal consonants (12%) than for initial consonants (20%). This lower proportion of voicing confusions for terminal consonants was found in all but one of the filtered conditions (1250 Hz) for the normal group, and in all but two clinical subjects (WB and SR).

Nasality errors were in general less frequent for the clinical group than for the normal group, especially on the terminal consonants.

The feature analysis of vowel errors for the clinical subjects is presented in Table 22. The number of vowel test items presented to each subject was 40. For most subjects the number of vowel errors was quite small. (Thus, an entry of 100% error for a particular feature may represent a single confusion.)

In general, place confusions are the most frequent, and duration errors are second most frequent. This order of feature confusion is the same as that for the experimental conditions. The most interesting vowel error finding is that none of the clinical subjects made a single multiple-feature confusion. Thus, these subjects never confused more than one vowel feature at a time. The normal-hearing listeners did make some multiple-feature vowel confusions, but these errors were always the least frequent for any filter condition in which they did occur.

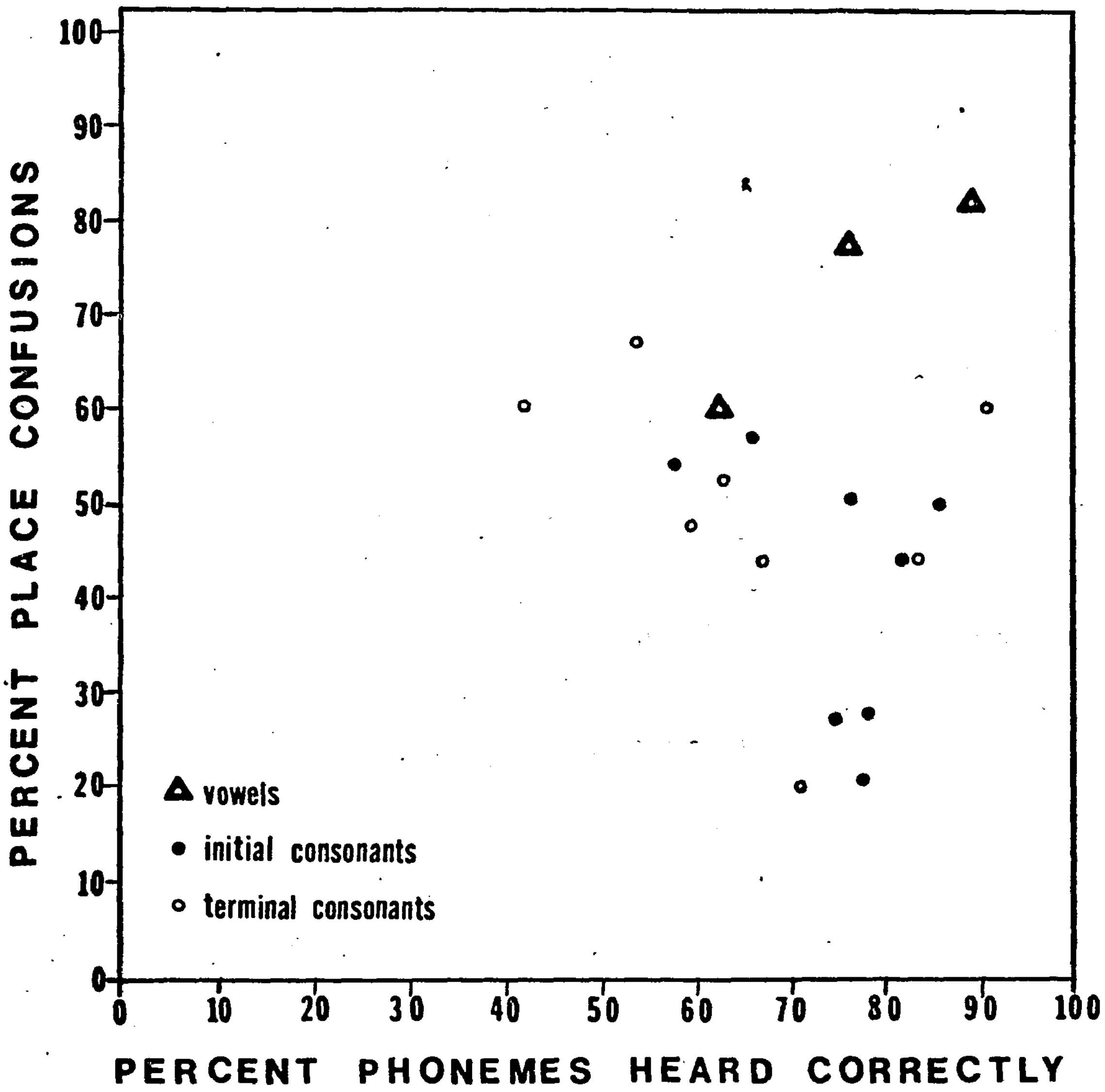
For the normal-hearing listeners under filtered speech conditions, a linear positive relationship was found between the proportion of place confusions and the intelligibility of both consonants and vowels. Figure 28 plots the proportion of place confusions against the percentages

TABLE 22

FEATURE ANALYSIS OF VOWEL ERRORS FOR NINE HARD-OF-HEARING SUBJECTS

Subject	Total Number of Vowel Errors	Percentage of Error Types			
		Place	Openness	Duration	Multiple
TO	2	50.0	0.0	50.0	0.0
EB	16	62.5	12.5	25.0	0.0
TE	9	77.7	0.0	22.3	0.0
CO	1	0.0	100.0	0.0	0.0
WB	5	100.0	0.0	0.0	0.0
JM	1	100.0	0.0	0.0	0.0
SR	0	---	---	---	---
JS	2	0.0	100.0	0.0	0.0
MH	2	0.0	50.0	50.0	0.0
All Subjects	38	64.0	15.0	21.0	0.0

Figure 28. Proportion of place confusions for initial consonants, terminal consonants, and vowels plotted against percent correct for nine hard-of-hearing subjects.


PERCENT PLACE CONFUSIONS
100
90
80
70
60
50
40
30
20
10
0
0
10
20
30
40
50
60
70
80
90
100
PERCENT PHONEMES HEARD CORRECTLY
▲ vowels
● initial consonants
○ terminal consonants

of consonants and vowels heard correctly by each of the clinical subjects. Only those points representing more than 5% error are plotted. The points represent the full scores for initial consonants, terminal consonants, or vowels. Further division into phoneme classes would involve extremely small numbers of errors for each point.

The plot shows considerable scatter, and for consonants there does not appear to be any linear relationship between intelligibility score and proportion of place confusions. This could be because of the small number of errors represented by each point. The question of whether this relationship differs for hypacusic and normal listeners can only be answered in a larger study.

Only three subjects made more than 5% vowel error. These three do, however, exhibit an increase of place confusions with increase of intelligibility, but one cannot establish a correlation for such a small sample.

The results for these nine hard-of-hearing subjects indicate that the PD test might be a useful clinical tool. All subjects were able to complete the test, and the scores for most subjects were reasonably similar to the scores obtained on earlier clinical discrimination tests. The greatest differences between PB and PD scores were found for the two youngest subjects, both of whom showed large differences between test and retest PB scores as well.

It appears that most subjects will score lower on the PD test than on a PB list. However, this is not because of any overall reduction of intelligibility for the PD list, since normal-hearing listeners achieved nearly perfect scores when the list was presented unfiltered. The PD test apparently involves finer discriminations than does an open-set PB list. Another factor to be considered here is the length of the PD list used in the present study. The full list is four times longer than the PB-50 list and may have produced fatigue, which might lower the performance. In future uses of the test, shorter lists could be used and might reduce the difference in scores between the PD and PB lists.

The general order of intelligibility of phoneme classes for the clinical group parallels the normal group, although certain trends may have been stronger or weaker than in the normal-hearing group.

Any differences found between the normal and clinical groups may result from the sample size or from individual differences within the clinical group. On the other hand, differences might result from distortions within the pathologic ear which are more complex than low-pass filtering.

Differences within the clinical group may be idiosyncratic or they could be related to the natures of the various hearing pathologies. For example, the listener with retrocochlear damage (WB) was the only subject to score higher on the PD test than on the PB list.

CHAPTER V

DISCUSSION

The Phoneme Differentiation Test described in this study was designed to provide various ways to analyze the efficiency of any type of speech transmission system. Like previous articulation tests, it furnishes a percentage score which estimates the overall intelligibility. But in addition, it gives more detailed information than open-response tests and affords more control over the alternative responses than previous closed-response tests.

Other closed-response tests (House _et_ _al_., Griffiths, Pickett _et_ _al_.) present sets of five or six response choices differing from each other in only one phoneme. The choices were generally designed to provide certain single or multiple feature contrasts. In these tests, any of the response words may be presented as the stimulus. However, the minimal features contrasted in the response alternatives when one word is the stimulus _are not_ and _cannot_ be the same features contrasted when a different response word is the stimulus. Given the test set "bail, mail, pail, veil, dale," minimal contrasts of nasality, voicing, manner, and place are available only when "bail" is the stimulus. When "mail" is the stimulus, no place contrast is available. This reduces the probability of error, since generally place confusions are the most frequent;

that is, the most frequent confusion of /m/ is with /n/. Given the same test set, when "pail" is the stimulus, the probability of error is also reduced because any alternate response would involve a voicing confusion, and this has been shown to be a very infrequent type of error. When "dale" is the stimulus, a place contrast, "bail," is available, but all of the other contrasts involve place as well.

In the Phoneme Differentiation Test, the choices for every stimulus word present contrasts in only one feature (within the phoneme and word inventory limitations of the English language). The set of words used in the above example appears on the test, but only when "bail" is the stimulus. Hence, the alternative responses provide the minimal contrasts of nasality in "mail," voicing in "pail," manner (openness) in "veil," and place in "dale." When "veil" is the stimulus, a different set of responses is presented so that again the alternative responses provide contrasts in only one feature: nasality in "mail," voicing in "fail," manner in "bail," and place in "they'll."

The major advantage of the minimal-contrast technique as employed in the PD Test is that is simplifies the analysis procedure. Each response is designated in advance as a confusion in a particular feature for a particular phoneme. With the other closed-response tests, the feature confusions must be determined through a rather complicated process of elimination, especially when the response differs from the stimulus in more than one feature.

Another advantage of the PD Test is that all minimal contrast responses are available in approximately equal proportions. Within a response set, all contrasts are provided once, assuring equal probability for each type of feature confusion. In addition, the probability of any type of error for each phoneme is equalized to the extent that all are provided with easy as well as difficult discriminations. Since place is a frequent confusion, the probability of error on a stimulus for which two place contrasts are available would be greater than on a stimulus for which no place contrasts are available. When the place feature is contrasted once for all phonemes, the probability of error is more equalized.

The closed-response format described above is essential for a valid analysis of the relative intelligibility of phonemes or classes of phonemes. Since the phoneme combinations which comprise words in a given language and the relative familiarity of words play a large role in the perception of speech, these factors must be accounted for when one tries to analyze phoneme substitutions. When the entire message set is known, as in the PD Test, the effects of word availability and word familiarity are minimized.

The closed-response technique allows testing of one phoneme at a time. With this format one can establish, for example, that vowels are more intelligible than consonants, given these test alternatives and conditions. When open-response tests are used, listeners can substitute entirely new words as responses; and in two analyses of such substitutions by hard-of-hearing listeners (Oyer and Doudna, 1959; Schultz, 1964), it appeared that vowels were less intelligible than consonants. However, when vowels and consonants are independently tested, each with the same probability of error, as in the present study, it is clear that vowels are generally more intelligible. This was found for all of the clinical subjects and under all but the extreme low-pass filter conditions. It is possible that certain hard-of-hearing individuals or pathologic groups may show a higher proportion of vowel errors, but the open-response tests cannot establish this.

In the present study, it was also shown that, for both clinical and experimental groups, consonants are perceived better in the initial word position than in the terminal word position. Pickett _et al_. (1972) reported a large advantage of vowels over consonants, and initial consonants over terminal consonants, for their group of 99 hypacusics. They used a closed-response test. However, even their most severely hard-of-hearing subjects did not reverse the order of vowels over consonants, as was found for the 250 and 500 Hz filter settings in the present experiment.[1]

In the present study, it has also been shown that some classes of phonemes are easier to discriminate than others.

---

[1] As suggested in Chapter IV, the phonemic feature information found in the frequencies below 500 or 250 Hz warrants further exploration. The pilot data showed that there is a certain amount of consonant information in that portion of the spectrum. These frequencies are of special interest when we consider the profoundly deaf who frequently respond to these low frequencies and not to any higher frequencies.

In both initial and terminal word positions, nasals were more intelligible than plosives, which were more intelligible than fricatives. Semivowels were the easiest to discriminate, and affricates were the most difficult. This hierarchy of intelligibility for phoneme classes cannot be validly established using open-response tests because the probability of error for each phoneme may not be equal, as explained above.

In addition to ranking the intelligibility of phoneme classes, one can discover the particular features responsible for the confusions. It has been shown in the present data that with sharp low-pass filtering, the place feature is very poorly transmitted for both consonants and vowels. Voicing and nasality are rather well preserved features for the consonants under these conditions. Miller and Nicely also reported that when high frequencies are removed, place errors predominate; but as more frequencies are removed, errors become more random in that other features are also frequently confused. The clinical data in the present study suggests that these statements also apply to hypacusics. Further support comes from the study of Pickett _et al_. They reported place is the least intelligible feature for both consonants and vowels. They also found that the subjects with the poorest hearing make more random confusions.

While the multiple-choice test used in the Pickett _et al_. study did not provide all one-feature contrasts, it is interesting to see the most frequent vowel substitutions they found: /u-i/(place); /i-I/(duration); /I-ʌ/(openness); /ɑ-æ/

(place); /ɛ-æ/(openness); /i-u/(place); and /ɑ-ɔ/(openness).
The features given in parentheses are those assigned to these confusions by the PD Test phoneme model. All of these are errors in one vowel feature.

Another finding in the present data which is also reported by Pickett _et al._ is that of fewer voicing errors on terminal consonants because of the longer vowel duration cue preceding voiced consonants.

Applications of the Phoneme Differentiation Test

The kind of feature analysis described above can be of value whenever one is testing an adjustable speech transmission system. A transmission system can be considered adjustable if its speech transmission efficiency can be changed in any way. Such changes in efficiency might result from adjustments in spectrum or intensity. The effects of such adjustments in spectrum or intensity on the transmission of the individual speech features can be determined by the feature analysis described here. Once a poorly-transmitted feature is known, it may be possible to improve transmission of that feature. If, however, this improvement comes at the expense of other more intelligible features, it might be more efficient to make greater use of the more intelligible features.

What is meant here by an "adjustable system" can be illustrated by a summary of the applications of the Phoneme Differentiation Test.

In the present study, the PD Test was used to assess the transmission of the phonemic features when low-pass filters were adjusted to remove the spectrum above various frequency points. There were several reasons for the selection of this particular application. First, low-pass filtering is a good example of an adjustable system. Second, there are data in the literature (Miller and Nicely) with which the present data could be compared. Third, low-pass filtering appears to be related to the most common type of hearing loss in that most losses are greater for the high frequencies, and as hearing losses become more severe, more frequencies become involved.

This analogy between hearing loss and low-pass filtering is probably not complete. Hypacusics may have speech discrimination that is poorer or better than would be expected from filtering data.[2]

However, the data obtained in the present study would be directly related to electronic speech transmissions systems with specifiable frequency limitations, e.g., receivers, microphones, amplifiers, and so forth.

Other obviously adjustable speech systems are electronic speech synthesizers, of which there are two basic types. One type of synthesizer begins with a real speech input, reconstitutes it, emphasizes or changes certain speech characteristics, and produces the synthetic speech signal. Examples of this type include the vocoder (Dudley, 1939) and

---

[2] For example, Rhodes (1966) demonstrated that a group of individuals with high frequency losses were better at perceiving filtered words than were a group of normal listeners.

the helium speech converter discussed below.

The purpose of the converter was to improve communication between a control center and divers working in a helium atmosphere in underwater tanks. The inhalation of helium changes the speaker's formant spectrum and the converter must transpose these distortions. The PD Test was used to estimate the distortion of phonemic features by the helium medium and the converter's ability to reverse the distortions (Flower, 1969).

The second type of synthesizer operates without a speech input, and truly synthesizes speech signals from electronically produced sounds according to rules based on phonemic and sub-phonemic features. A test of these rules should provide more than a score or a tabulation of unsatisfactory stimuli. For a given phoneme or class of phonemes, it would be more efficient to test for the troublesome feature. In this way, one would know the particular rule, or rules, needing adjustment. The PD Test was used by Haggard and Mattingly (1968) for this purpose. PD Test items were synthesized according to Haggard and Mattingly's rules for Southern British pronunciation, and these stimuli were presented to speakers of that dialect,using the closed-response technique. Their erroneous responses indicated the feature, and thus the rule, which needed adjustment in order to produce a more satisfactory synthetic phoneme.

In clinical settings we also have adjustable speech transmission systems. The hearing aid is the most obvious example, but the listener,himself, is adjustable. First, his speech discrimination may be diminished with onset and progression of various otopathologies or central disorders. Second, his speech discrimination may be improved through experience and training. The PD Test would provide an estimate of the auditory cues being used by the hypacusic listener prior to training, outline the cues to be taught, and later, test the effectiveness of the training program in the improvement of speech discrimination by the listener. In addition, if groups of listeners with hearing losses of long duration and listeners with comparable hearing losses of recent onset could be tested, the differences between the groups would indicate the cues which can be learned and used as compensatory devices in speech perception. If these cues were known, perhaps auditory training could further hasten the learning processes. The PD Test format might even be utilized as a training technique for this purpose.

One must also consider the adjustability of the speaker's voice. The PD Test is currently being used in a study of the effects of voice types--male, female, whispered, and dysphonic--on the perception of phonemic features by hard-of-hearing listeners.[3]

Modifications of the Phoneme Differentiation Test

For general clinical and experimental uses, the PD

---

[3] Kirsch, Edward. M.A. thesis in progress, Hunter College of the City University of New York.

Test can be divided into four word-lists, each of which tests the same fifty phonemes. The number of errors for each of the 200 test words has been tabulated to determine the four fifty-item combinations which yielded equivalent scores in the main study. The individual word scores are presented in Appendix B, and the equivalent lists are presented in Appendix C. These lists are of appropriate length for clinical uses. Additional lists might be developed by adding new test items according to the feature model, or by randomizing items across the four lists.

Other modifications of the PD Test may be required for special clinical uses. The vocabulary used in the present word-list is probably too advanced for children, or for those who have recently learned English. However, a modified version can be developed employing the same minimal-contrast criterion, though the test phoneme inventory may be limited if the vocabulary is restricted.

In certain test situations, nonsense syllables may be the necessary or preferable type of stimuli. When trained listening crews are employed for repeated testing of electro-acoustic systems, limited lists of actual words may be learned and, therefore, be invalid as test stimuli. However, a great number of different lists of nonsense syllables can be derived. If phoneme or feature confusions are of interest in systems testing, syllable lists can be derived according to the PD Test minimal-contrast technique.

Other studies may also require the use of nonsense syllables as stimuli to avoid learning of the lists. In studies of short-term memory, lists of actual words are inappropriate as stimuli because of the semantic factor which would influence memory. One could present lists of nonsense syllables for immediate recall. At the end of the presentation, listeners would respond by selecting one of five nonsense syllable choices for each stimulus presented. This would avoid the unequal probabilities of error confusions and response bias encountered in a procedure such as Wickelgren's (1965a, 1965b). In his studies, listeners selected any of sixteen or twenty-three consonants, or any of six vowels. For consonant stimuli, there are usually several place contrasts available, but only one voicing or nasality contrast. Using the minimal-contrast technique, each feature contrast is available only once for each response. The features shared by different phonemes within a set of response choices would be indicated by the confusions within the set. For example, assume that the test item presented on a given list of nonsense syllables is /tʃa/. Given the alternatives /tʃa, ka, ʃa, na, dʒa/, if the most frequent error response is ka/, this would suggest greater affinity of affricates for stops than for fricatives. The best of several feature systems could be chosen by systematically applying each system to a separate test, and comparing their relative success in predicting confusions. As a check on each system, multiple-feature contrasts might be included as response choices. The best feature systems would show a minimum of multiple-contrast errors.

The Selection of Minimal-Contrast Features

Feature analysis of errors in short-term memory (Wickelgren), or in perception (Miller and Nicely), has led to the conclusion that phoneme errors are not random. The substituted phoneme has more features in common with the stimulus than would be expected by chance. In other words, listeners confuse one feature independently of all other features. Support for this theory has been indicated at several points in the preceding chapter of this paper. Listeners rarely gave responses differing from the stimulus in more than one feature. Although the PD Test was designed to avoid multi-feature contrasts for the consonants, because of the constraints of the English language phoneme inventory, several were present. For example, since there are no voiceless nasals or voiceless semivowels, voicing contrasts to these phonemes necessarily involved voicing and openness simultaneously. However, such voicing confusions were extremely rare for nasals and semivowels. This suggests that listeners did not confuse more than one feature at a time. In addition, the multiple-feature contrast provided for the vowels was rarely chosen.

It must be remembered that any conclusions regarding the predominance of single-feature errors in any set of data are highly dependent on how the features are defined and applied. If one accepts the theory which states that phonemes are combinations of features and that each feature may be perceived somewhat independently, if follows that the feature model which predicts the greatest number of confusions in any set of data, is the feature model most representative of perceptual processes. In other words, the ways in which we confuse features relate to the ways in which we perceive features. No feature system appears to be completely accurate in predicting these confusions.

There are two possible approaches to the task of developing improved feature systems. The first approach is to define a set of features based on previous studies or theoretical considerations. The accuracy of the features is then tested by determining the number of correct error predictions for a given set of data. If phonemes which have a feature in common are frequently confused with each other, that feature would appear to be valid. If, however, all phonemes which share a feature do not group together, one concludes that the feature is not valid, or, that while the feature may be valid, not all of the phonemes were correctly described with regard to this feature.

This approach, namely, <u>evaluation</u> of predetermined sets of features, has been the most commonly used. It was employed by Miller and Nicely, Wickelgren, Studebaker, Pickett <u>et al</u>., and others mentioned in this paper.[4] This analysis can also be applied to the present data. A discussion of feature evaluation in the present study appears below.

---

[4] It should be noted here that no two studies evaluated the same set of features.

The second, and more promising procedure for developing an improved feature system is the derivation of features from a given set of phoneme confusion data, prior to any theoretical considerations. With this procedure, the first step is to group the phonemes which are most frequently confused with one another. These phoneme groups are then examined to decide if any features can explain the groupings. For example, if one group of ten phonemes included nine fricatives, one might conclude that frication is a feature.

This approach was introduced by Klatt (1968). Using a similarity metric, Klatt estimated the presence in Wickelgren's consonant data (1965b) of each of the binary features of Chomsky and Halle (1968). Klatt found that the features voiced, long frication, sonorant, continuant, and strident were more strongly indicated than anterior, coronal, nasal, and consonantal. However, the actual indications for certain features were reduced when their dependencies upon other features were evaluated. Stridency, for example, appeared to be a redundancy of long frication. Klatt also found that a multi-dimensional place feature had a high score, and this explained the weakness of the binary place features anterior and coronal.

In a further analysis of Wickelgren's data, Klatt derived optimum features by grouping the phonemes which were confused with each other without consideration of any of the linguistic features. The optimum features which emerged were not those to be expected from linguistic feature descriptions, but the first feature derived did group the fricatives together, allowing some confidence in the technique. While the optimum feature analysis does not verify any one feature system, it does provide an objective evaluation of feature systems and should be applied to other data.

The derived feature technique could not be applied to the present data because a predetermined set of features was used to structure the test sets. Confusions were limited to the phonemes presented as choices, and the choices were those I considered to be feature contrasts.

Klatt's technique could be applied to perceptual confusion data, provided the choices are not limited to only certain phonemes. Wickelgren's memory confusion data was suitable for feature derivation because all consonants were available as responses. Feature derivation from the data of Miller and Nicely would not be possible because responses were limited to only sixteen consonants. Klatt's procedure should be applied to perceptual confusion data obtained when all consonants are available as responses. The features thus <u>derived</u> could then be <u>evaluated</u>, using the minimal-contrast technique as employed in the PD Test.

Until the best possible features are derived, one must utilize the features suggested by previous data. In the present data, evaluation of features is possible to a limited extent. The low-pass filtering effects on the stimuli probably are responsible for the relative intelligibility of each of the features tested. Since place information is found in the high frequency bursts and transitions (Delattre <u>et</u> <u>al</u>., 1955), a high proportion of place errors would be expected in these data. Openness confusions would be less frequent, because the cues to openness involve lower frequencies and may be dependent on the rates of transitional changes (Liberman <u>et</u> <u>al</u>., 1956). Therefore, openness information is less affected by frequency distortion. We would also expect voicing information to be preserved in these stimuli, because voicing cues are found in the low frequencies and are also dependent on timing, rather than on spectrum (Lisker and Abramson, 1964, 1967).

Based on the assumption that phonemes frequently confused with each other are alike in all but one feature, place is strongly present as a feature in the present data, since it is responsible for the majority of errors.

Duration, as employed by Miller and Nicely, does not appear to be a consonant feature in the present data. Duration was Miller and Nicely's feature adopted to distinguish /s,z,ʃ,ʒ/ from the other consonants. They found that when frequencies above 1000 Hz were removed, the place and duration distinctions were lost. It seems more likely that these were all actually place confusions. Yet, the cross-over point of equal intelligibility was 2200 Hz for the duration feature, and 1900 Hz for the place feature. The higher cross-over point which they found for the duration feature is probably because /s,z,ʃ,ʒ/ are very high frequency consonants. In the present data, /s,z,ʃ/ appear to operate in the same manner as the other fricatives. There is no evidence in these data for a feature which separates phonemes within the class of fricatives. The phonemes /s,z,ʃ/[5] were confused as frequently with /θ,ð,f,v,h/ as they were with each other.

As stated above, the high proportion of place confusions suggests that place of articulation is a feature as defined in the PD feature model. However, the low proportions of confusions involving the other features does not mean that they are not valid features. As already explained, the features confused were dependent on the spectra of the stimuli presented. If different conditions of distortions produced a high proportion of errors involving some other feature as defined in the model, one could then conclude that such a feature was valid.

Feature Descriptions of Individual Phonemes

The present data may be useful in evaluating the descriptions of the individual phonemes in terms of the features tested. That is, assuming openness is a feature, the question arises: how should each phoneme be described according to this feature? For example, how "open" is the /h/? Chomsky and Halle grouped /h/ with /w,r,l,j/ as sonorant, and with /w,j/ as non-consonantal. Wickelgren assigned /h/, as well as /w,r,l,j/, a greater degree of openness than

---

[5] The phoneme /ʒ/ was not tested, because suitable test sets could not be found.

the fricatives, which in turn, were assigned a greater degree of openness than the stops. These classifications of /h/ lead to the expectation of more confusions with semivowels than with fricatives or stops. However, any confusion of /h/ with a semivowel would also involve voicing and therefore be unlikely. In the present data, /h/ was most frequently confused with /k/, unless /ʃ/ was available as a choice. Confusions of /h/ with /s/ or /θ/ were less frequent than confusions of /h/ with /k/ or /ʃ/. This casts doubt on the high degree of openness assigned to /h/ by Wickelgren. If /h/ were more open than other fricatives, the confusions of /h/ with /k/ would involve two-step confusions (/h/ having an openness value of 2, and /k/ having an openness value of 0). These should be less frequent than the one-step confusion of /h/ with fricatives (openness value of 1).

In the PD feature model, the nasals were classified with the stops, having an openness value of 0. Nasality contrasts were presented in voiced stops, while openness contrasts were presented in semivowels or voiced fricatives. Both the contrasts actually involve differences in nasality. If nasals are more like stops than fricatives or semivowels, there should be a greater number of confusions with stops (nasality confusions) than with fricatives or semivowels (openness confusions). In the present data, there were twenty-five "openness" confusions for /m/ and /n/, but only four "nasality" confusions. Thus, the nasals are perceived as continuants far more frequently than they are perceived as stops. This suggests that the phoneme model should have classified the nasals with the semivowels or fricatives, and not with the stops, as it did. Considering the frequent semivowel-for-nasal substitutions found in the present data, and the vowel-like qualities of the nasal consonants, it would probably be better to classify the nasals with the semivowels in terms of the openness feature.

The openness classification of the affricates with stops, rather than with fricatives, was discussed in Chapter IV. The data indicated that /tʃ/ and /dʒ/ are confused with stops more frequently than with fricatives, and, thus, should be classified as stops.

Certain other questions regarding phoneme descriptions cannot be answered from the present data, because not all contrasts were available for comparison, and because only one feature system was tested.

Comparisons of the intelligiblity of consonants and vowels have been made in this study as well as in previous studies. These direct comparisons are not entirely valid, however, because the features for vowels and consonants are not the same in kind or in number.

Further exploration with test sets designed to establish better feature descriptions might be of value. The purpose of the present study was to develop a useful technique for confusion analyses based on feature information already available, but the technique itself could be used to test other feature systems.

CHAPTER VI

SUMMARY

A new CVC multiple-choice articulation test was developed for use in phoneme and feature confusion analyses. The test may be applied to all types of speech transmission systems. The Phoneme Differentiation (PD) Test is similar to other recently devised closed-response tests, in that, for each set of choices, the stimulus varies from the available responses in only one phoneme. However, in the PD Test, syllable nuclei, initial consonants, and final consonants are each tested with the same format and, therefore, the same probability of error.

The major difference between the Phoneme Differentiation Test and previous closed-response tests is the structure of the response choices. For each stimulus word, the alternatives were specially tailored to provide contrasts in only one phonemic feature at a time. Thus, the set of contrasts for each stimulus is different from the set of contrasts for every other stimulus. In this way, the four consonant features--voicing, nasality, manner of articulation, and place of articulation--are contrasted for every consonant stimulus. The vowel features--intrinsic duration, openness of the vocal tract, and place of articulation--are contrasted for every vowel stimulus. The fourth vowel contrast involves two or more features simultaneously, and can be considered the relatively easy discrimination for vowels, just as the nasality feature can be considered the relatively easy discrimination for initial consonants, and the voicing feature can be considered the easy discrimination for final consonants.

The PD Test lists consists of 200 test items distributed as follows: four tokens each of 22 initial consonant contrasts, 13 final consonant contrasts, 10 vowel contrasts, and 5 diphthong contrasts. The full list has been divided into four 50-item sub-lists, each of which tests the same phonemes.

The major advantages of the test are the ease and uniformity with which feature confusions can be tabulated, and the assurance that all feature confusions are available as response choices.

The test was applied in a low-pass filtering experiment, and the proportions of each feature confusion under each filter condition were presented and discussed. The test was also given to a sample of hypacusic listeners so that the clinical feasibility of the procedure could be determined.

The results for the 104 normal-hearing subjects who heard the list under various filter conditions, and the nine clinical subjects who heard the list unfiltered, may be summarized as follows:

1. When the unfiltered Phoneme Differentiation Test

was presented to normal-hearing listeners at 56 dB SPL, nearly perfect scores were obtained. The mean score for the twelve listeners was 98%. This means that when lower scores are obtained, they are not because of variables within the test list or the recordings used.

2. The data obtained from normal-hearing listeners under several low-pass frequency conditions showed that intelligibility was proportionally reduced as a function of lower cut-off frequency.

3. For the normal-hearing listeners, under all but the poorest listening condition (all frequencies above 500 Hz removed), and for the group of hypacusics, vowels were perceived better than consonants. Further, consonants in the word-initial position were perceived better than consonants in the word-terminal position.

4. For all experimental subjects and most clinical subjects, place of articulation was clearly the most frequently confused feature for consonants and vowels. For consonants, openness (manner) was the second-most frequently confused feature, while the voicing and nasality features were infrequently confused. For vowels, openness was the second most frequently confused feature under some conditions, while duration was the second-most frequently confused feature for other conditions.

5. Multiple-feature confusions provided for all vowel stimuli were very infrequent, and showed a clear decline as the listening condition improved. In the clinical

group, there was not a single multiple-feature vowel confusion. Those multi-feature contrasts available for some consonants (because of the phoneme limitations in the language) were rarely, if ever, chosen.

6. In both experimental and clinical groups, voicing errors were less frequent on consonants in the final word position than in the initial word position, suggesting that both groups of listeners utilize the cue of longer vowel duration preceding voiced consonants.

7. With a low-pass setting of 250 or 500 Hz, the vowels lose their superiority over consonants. At the 250 Hz setting, vowel scores were near chance level, while consonant scores were well above chance.

8. The lower proportion of place confusions found for vowels as compared to consonants is due to the asymmetry of the vowel chart. Vowels change in the openness feature as they change in the place feature. If the vowel model is strictly applied, many of the needed minimal-contrasts do not exist. For the three vowel features used and the ten vowels tested, only 19 of the 30 needed minimal-contrasts were available. The fact that the features applied to the vowels were less symmetrical and fewer in number than were the features applied to the consonants may partly explain the higher vowel scores obtained. Since many of the vowel contrasts involved more than one feature simultaneously, the probability of error may be lower for the vowels than for the consonants which were almost always contrasted with

phonemes differing in only one feature.

9. The proportion of place errors for individual classes of both consonants and vowels appears to be related more closely to the overall intelligibility of a phoneme class than to the filter setting.

10. With a less-steep filter slope (30 dB per octave as compared to 60 dB), there are slight increases in intelligibility, but no differences in the types of errors found.

11. The findings for both clinical and experimental groups are in general agreement with those of previous clinical and experimental studies.

12. Several of the findings in the present study lend support to the theory that phonemes are confused one feature at a time. In addition, a multi-valued place feature is strongly indicated.

Several applications of the PD Test and the minimal contrast technique it employs have been suggested. In its present form, it could be used as a clinical test of the diagnostic significance of feature errors, and as a means of summarizing the rehabilitative needs of hypacusic listeners. It may also be useful in the testing of electronic speech transmission systems, such as speech-input synthesizers, rule-based synthesizers, and other speech transmission equipment.

The technique could be adapted to nonsense syllables for use in studies which utilize trained listening crews in the evaluation of speech transmission systems.

APPENDIX A

PD TEST ORDER 1-A:
RESPONSE FORM
ANSWER KEY

NAME ______________________

AGE ______________________

DATE ______________________

TEST LOCATION ______________

TEST CONDITION _____________

EAR TESTED _________________

INSTRUCTIONS

You have been given answer sheets containing rows of five (5) words each. You will hear one word from each row. You are to circle the word that you hear. Do not circle more than one word in each row. If you are not sure, GUESS! No item is to be left unanswered. There will be a short pause after every five rows.

- B E G I N -

COLUMN 1

GAVE	SHAVE	CAVE	SAVE	KNAVE
MAY	BAY	WAY	RAY	LAY
WIDE	LIED	DIED	RIDE	BIDE
FAIR	MARE	TEAR	CARE	PAIR
RARE	DARE	THERE	MARE	FARE

SOAK	CHOKE	COKE	JOKE	POKE
PART	MART	DART	TART	BART
LORE	BORE	CORE	GORE	DOOR
CARD	HARD	GUARD	TARRED	YARD
VEND	BEND	PEND	FEND	MEND

MIT	MEET	MAT	MATE	MUTT
FOUGHT	FIT	FAT	FEET	FOOT
LED	LOAD	LAD	LID	LEAD
MID	MADE	MAD	MOOD	MODE
FILL	FOOL	FEEL	FELL	FILE

BAT	BASS	BASH	BATH	BAN
GAZE	GAVE	GAME	GAIN	GATE
PIN	PITCH	PIT	PICK	PIG
RICK	RICH	RIP	RIDGE	WRIT
SPIED	SPINE	SPICE	SPIES	SPITE

COLUMN 2

COT	NOT	DOT	WATT	GOT
SALT	MALT	FAULT	PALT	VAULT
PINK	THINK	SINK	ZINC	MINK
NEST	TEST	JEST	CHEST	FEST
DEED	BEAD	NEED	TEED	LEAD
DO	TO	GOO	NEW	ZOO
TILL	HILL	KILL	NIL	GILL
MAT	FAT	BAT	VAT	THAT
BAIL	MAIL	PAIL	VEIL	NAIL
MY	THY	BY	VIE	THIGH
THY	TIE	THIGH	SHY	MY
JET	YET	CHET	BET	NET
USE	NEWS	SUES	LOSE	ZOOS
DO	NEW	ZOO	MOO	TOO
VOLE	POLE	MOLE	BOWL	GOAL
GOAL	HOLE	KNOLL	TOLL	COAL
POUR	WORE	MORE	BORE	NOR
THINE	FINE	MINE	DINE	VINE
POT	NOT	TOT	SOT	DOT
RIND	MIND	FIND	BIND	WIND

COLUMN 3

MORE	PORE	FORE	BORE	DOOR
JAIL	NAIL	WAIL	YALE	SHALE
FAT	HAT	NAT	THAT	CAT
HOT	NOT	COT	GOT	POT
VEIL	MAIL	BAIL	FAIL	THEY'LL
PILL	FILL	TILL	BILL	MILL
TIE	BUY	THY	NIGH	DIE
LEER	DEAR	NEAR	MERE	TIER
GOAL	HOLE	KNOLL	COAL	SOLE
SHOE	YOU	WOO	NEW	GOO
ROLL	BOWL	KNOLL	GOAL	COAL
JOAN	SOWN	LOAN	KNOWN	ZONE
SHIP	SIP	NIP	KIP	ZIP
VIC	PICK	THICK	NICK	SICK
MAN	BAN	TAN	PAN	FAN
SHEET	HEAT	WHEAT	NEAT	KEET
PAT	CAT	MAT	FAT	BAT
NIL	KILL	DILL	SILL	TILL
SEAL	VEAL	MEAL	FEEL	PEAL
FORE	ROAR	WORE	BORE	MORE

COLUMN 4

PAT	CAT	NAT	GAT*	HAT
BAR	CZAR	CHAR	MAR	JAR
CHILL	KILL	JILL	SHILL	NIL
WILL	BILL	GILL	NIL	KILL
VET	PET	GET	BET	MET

SORE	NOR	TORE	DOOR	PORE
DO	ZOO	VOO	NEW	SUE
NEWS	SHOES	ZOOS	WOOS	RUES
GALE	KALE	NAIL	DALE	WAIL
YACHT	NOT	SHOT	GOT	LOT

LIP	SIP	NIP	RIP	GYP
DUDE	RUDE	NUDE	CHEWED	JUDE
WAIT	SATE	RATE	MATE	DATE
KNOCK	LOCK	MOCK	TOCK	DOCK
SHEER	FEAR	PIER	MERE	VEER

BAT	THAT	MAT	VAT	FAT
REAR	JEER	YEAR	SHEER	NEAR
CHEWED	COOED	JUDE	MOOD	SUED
SUES	TWOS	NEWS	ZOOS	SHOES
ZONE	KNOWN	SHOWN	CONE	PHONE

COLUMN 5

TIP	NIP	SIP	DIP	KIP
KNOWN	ZONE	TONE	SHOWN	SOWN
DOT	NOT	LOT	SOT	ROT
DEEP	LEAP	NEAP	SEEP	WEEP
CHAR	FAR	TAR	MAR	JAR

FINE	VINE	BINE	MINE	THINE
VEAL	DEAL*	KNEEL	SEAL	ZEAL
NAB	GAB	DAB	LAB	TAB
WENT	LENT	MEANT	BENT	SENT
TAUGHT	FOUGHT	BOUGHT	SOUGHT	NOUGHT

VEIL	PAIL	MAIL	BAIL	DALE
WILL	MILL	PILL	BILL	NIL
YALE	DALE	NAIL*	TAIL	MAIL
DEN*	ZEN	MEN	HEN	THEN
NIP	DIP	CHIP	GYP	RIP

KIN	SHIN	MIN	GIN	FIN
SAD	MAD	LAD	TAD	SHAD
DEAD	NED	SAID	JED	ZED
FALL	BALL	PAUL	CALL	MALL
BUT	NUT	RUT	MUTT	PUTT

COLUMN 6

SHINE	FINE	VINE	MINE	PINE
ROT	NOT	JOT	YACHT	SHOT
MINK	CHINK	FINK	SINK	ZINC
COME	RUM	NUMB	HUM	THUMB
SIP	LIP	NIP	RIP	DIP

FAIL	BAIL	WAIL	YALE	MAIL
VEST	NEST	BEST	ZEST	FEST
BOUGHT	NOUGHT	TAUGHT	SOUGHT	THOUGHT

COLUMN 7

SOT	SIT	SUIT	SAT	SOOT
BEET	BOUT	BOAT	BAIT	BITE
FOUND	PHONED	FIND	FIEND	FEIGNED
PUP	PEEP	PEP	PIP	POP
POT	PAT	PEAT	PUT	PET

SUP	SEEP	SOP	SAP	SOUP
GATE	GERT	GOAT	GOT	GOUT
COD	COULD	KID	CAD	COOED
PEAK	PERK	PUCK	PECK	PICK
HACK	HAWK	HOCK	HICK	HOOK

BIT	BOOT	BOUGHT	BAT	BUT
PIN	PAWN	PUN	PEN	PAN
LUKE	LEAK	LICK	LOOK	LACK
BOYS	BAYS	BUYS	BEAUS	BOUGHS
PUT	PAT	PEAT	PIT	POT

ROY'S	ROSE	RAISE	ROUSE	RISE
SEWS	SOWS	SIZE	SAWS	SEES
KNIT	NUT	NET	NOT	NEAT
TOD	TEED	TIDE	TOYED	TOWED
POISE	PIES	POSE	PAYS	PAUSE

COLUMN 8

BOOT	BAIT	BOAT	BOUT	BITE
DONE	DAWN	DEN	DUNE	DEAN
BEAD	BUD	BIRD	BID	BAD
DICK	DUKE	DECK	DOCK	DUCK
BUG	BIG	BAG	BERG	BOG

TIM	TOM	TOMB	TAM	TUM
TALE	TILE	TOWEL	TOIL	TOLL
BOOZE	BEAUS	BOUGHS	BUYS	BOYS
KIN	KERN	KEN	KEEN	CON
PUT	POT	PAT	PEAT	PIT

KED	CUD	COULD	CURD	KEYED
COOP	KEEP	KIP	CUP	COP
CAP	CUP	COOP	COP	KIP
NAN	NOUN	NINE	NOON	KNOWN
PAT	PERT	PUTT	POT	PIT

LUKE	LOCK	LOOK	LICK	LEAK
CHICK	CHECK	CHOCK	CHUCK	CHEEK
DUNE	DAN	DONE	DIN	DAWN
TOMB	TAM	TIM	TEAM	TOM
BOUGHT	BOOT	BEET	BET	BUT

COLUMN 9

TALK	TACK	TICK	TUCK	TOCK
CAP	COOP	KIP	KEEP	COP
POSE	POISE	PAYS	PEAS	PIES
BEG	BERG	BOG	BIG	BUG
KEET	CAT	CUT	COOT	COT

SUP	SEEP	SOUP	SAP	SIP
BED	BID	BAD	BEAD	BUD
TAD	TOAD	TEED	TIDE	TOYED
HOT	HEAT	HURT	HOOT	HUT
COULD	CAWED	KID	CAD	COOED

LUKE	LACK	LUCK	LOCK	LICK
LOUD	LOAD	LIED	LLOYD	LEWD
TOM	TOMB	TERM	TEAM	TIM
TOWN	TONE	TEEN	TINE	TURN
BOOT	BOAT	BITE	BAIT	BOUT

LUKE	LUCK	LOCK	LURK	LEAK
ROY'S	RAISE	RISE	ROUSE	ROSE
LLOYD	LEAD	LOUD	LOAD	LIED
BET	BUT	BOOT	BERT	BEET
POSE	PIES	POISE	PEAS	PAYS

COLUMN 10

LIME	LIFE	LIVE	LIED	LITHE
ROOM	ROOF	ROOT	RUTH	ROOVE
TEED	TEASE	TEAM	TEETHE	TEETH
ROBE	ROAN	ROSE	WROTE	ROAD
BUZZ	BUS	BUN	BUD	BOVE
LOAVE	LOBE	LOAF	LOATHE	LOAN
CAME	CAGE	CAPE	CAVE	CANE
PUCK	PUFF	PUB	PUN	PUP
WING	WICK	WIG	WISH	WIT
KID	KIN	KIT	KISS	KICK
LINE	LICE	LIED	LIES	LIVE
RICK	RIM	RIP	RIFF	RIB
KING	KIN	KID	KIT	KILL
LAB	LALL	LACK	LAMB	LAG
LICE	LIFE	LIES	LINE	LIGHT
SID	SIS	SIN	SIP	SIT
BUT	BUZZ	BUN	BUD	BUG
REAP	WREATH	REAVE	REEF	REAM
RIP	RIB	RILL	RIG	RIM
BUZZ	BUNG	BUN	BUT	BUD

COLUMN 11

SEED	SEEN	CEASE	SEIZE	SEETHE
ROAM	ROPE	ROVE	ROAN	ROBE
BUN	BUT	BUS	BUZZ	BUFF
RIM	RID	WRIT	RIG	RILL
LEAN	LEAP	LEAVE	LEASH	LEAF
DOLE	DOSE	DOME	DOZE	DOGE
BUD	BUG	BUCK	BUNG	BUZZ
COB	CON	COL	COD	COP
LIED	LINE	LIGHT	LIKE	LICE
WELL	WHEN	WET	WEB	WED
DUCK	DONE	DOES	DUB	DUG
LEAK	LEAN	LEASH	LEAGUE	LEAP
MATE	MAZE	MAIN	MADE	MAIM
LEAN	LEASE	LEAK	LEASH	LEES
LEAD	LEAGUE	LEAN	LEAK	LEES
COME	CUFF	CUT	CUP	CUB
CODE	COKE	COACH	COPE	CONE
LOSE	LOOM	LOOT	LOOF	LOOSE
DUMB	DUB	DUFF	DOVE	DOES
RILL	RING	RIM	RIP	RIB

COLUMN 12

ROPE	ROVE	RODE	ROAM	ROBE
COAL	COAT	CODE	CONE	COPE
RODE	WROTE	ROSE	ROAN	ROAM
LOAF	LOBE	LOATHE	LOAM	LOAVE
LIME	LIED	LIFE	LIVE	LITHE

SOON	SUED	SUES	SOOTHE	SOOTH
BUS	BUCK	BUNG	BUG	BUT
LIFE	LIME	LIVE	LICE	LIGHT
SUP	SUNG	SUB	SULL	SUM
DUCK	DUB	DUG	DOVE	DUMB

WANE	WAIF	WAYS	WADE	WAVE
ROAM	ROACH	ROPE	WROTE	ROBE

ANSWER KEY TO TEST ORDER 1-A

COL.2	COL.3	COL.4	COL.5	COL. 6	COL. 7
got	bore	cat	tip	fine	soot
fault	Yale	jar	sown	rot	bait
think	hat	chill	lot	sink	found
chest	cot	gill	leap	hum	pep
deed	veil	bet	char	lip	pat
do	pill	tore	vine	wail	sap
kill	die	zoo	zeal	vest	gate
that	near	rues	dab	thought	could
mail	hole	gale	went		peck
Thy	you	yacht	sought		hook
thigh	goal	rip	bail		bought
jet	zone	Jude	mill		pun
lose	ship	rate	nail		leak
new	thick	knock	then		beaus
bowl	pan	fear	gyp		put
coal	heat	vat	shin		rise
more	pat	year	sad		sows
thine	till	chewed	zed		net
tot	feel	shoes	Paul		toyed
wind	war	shown	mutt		pies

COL. 8	COL. 9	COL.10	COL.11	COL.12
boat	talk	live	seize	robe
dune	keep	roof	roam	coat
bird	pays	teethe	bus	roan
duck	big	road	rid	loave
bag	cat	buzz	leaf	lithe
Tom	sip	loathe	doze	soothe
tale	bid	came	bug	buck
boys	tide	pup	cob	life
kin	hurt	wick	light	some
peat	cawed	kit	wed	dub
cud	lock	lies	dug	wave
coop	loud	rip	leak	rope
cop	term	kin	main	
nine	tone	lag	lease	
pot	bout	lice	league	
Luke	lurk	sit	cup	
check	rose	bud	coke	
dawn	Lloyd	reef	loose	
team	but	rib	dove	
boot	poise	bun	rim	

APPENDIX B

ORDER OF INTELLIGIBILITY OF PD TEST STIMULI:

VOWELS AND DIPHTHONGS

INITIAL CONSONANTS

TERMINAL CONSONANTS

VOWEL STIMULI

Rank	Phoneme	Individual Tokens and Numbers of Errors				Totals
1.	/ɪ/	BID 5	BIG 7	SIP 7	KIN 24	43
2.	/ɔ/	CAWED 4	BOUGHT 9	TALK 15	DAWN 16	44
3.	/u/	LUKE 8	COOP 13	BOOT 14	DUNE 16	51
4.	/ɝ/	LURK 11	BIRD 16	TERM 24	HURT 30	81
5.	/ɑ/	TOM 15	LOCK 20	POT 23	COP 28	86
6.	/ʌ/	CUD 18	PUN 23	DUCK 26	BUT 28	95
7.	/ʊ/	HOOK 14	COULD 17	PUT 21	SOOT 46	98
8.	/æ/	BAG 23	CAT 23	SAP 23	PAT 35	104
9.	/i/	PEAT 7	LEAK 14	KEEP 21	TEAM 69	111
10.	/ɛ/	CHECK 19	PECK 30	PEP 33	NET 35	117

DIPHTHONG STIMULI

Rank	Phoneme	Individual Tokens and Numbers of Errors				Totals
1.	/ɔɪ/	POISE 6	LLOYD 7	TOYED 19	BOYS 22	54
2.	/eɪ/	BAIT 9	GATE 13	PAYS 14	TALE 23	59
3.	/aʊ/	LOUD 21	FOUND 22	BOUT 23	SOWS 29	95
4.	/aɪ/	PIES 17	TIDE 25	RISE 26	NINE 30	98
5.	/oʊ/	ROSE 15	BOAT 23	TONE 24	BEAUS 38	100

INITIAL CONSONANT STIMULI

Rank	Phoneme	Individual Tokens and Numbers of Errors								Totals
1.	/l/	LIP	0	LEAP	4	LOT	4	LOSE	15	23
2.	/w/	WORE	6	WAIL	8	WENT	16	WIND	20	50
3.	/r/	ROT	7	RIP	15	RUES	15	RATE	15	52
4.	/m/	MORE	2	MUTT	8	MAIL	21	MILL	27	58
5.	/h/	HOLE	11	HAT	11	HEAT	17	HUM	23	62
6.	/n/	NEW	15	KNOCK	17	NAIL	18	NEAR	20	70
7.	/p/	PAUL	12	PILL	17	PAT	19	PAN	28	76
8.	/g/	GOAL	13	GILL	20	GALE	20	GOT	33	86
9.	/f/	FEAR	12	FINE	20	FEEL	25	FAULT	34	91
10.	/j/	YACHT	17	YEAR	22	YOU	26	YALE	31	96
10.	/k/	COT	12	COAL	14	CAT	34	KILL	36	96
11.	/d/	DAB	16	DO	16	DEED	31	DIE	42	105
12.	/z/	ZONE	14	ZED	19	ZOO	23	ZEAL	50	106
13.	/s/	SAD	13	SOUGHT	14	SOWN	38	SINK	42	107
14.	/b/	BOWL	22	BAIL	24	BORE	38	BET	40	124
15.	/ð/	THEN	20	THINE	23	THY	35	THAT	51	129
16.	/v/	VINE	19	VEST	27	VEIL	39	VAT	55	140
17.	/θ/	THIGH	10	THINK	38	THICK	43	THOUGHT	55	146
18.	/tʃ/	CHEST	22	CHAR	27	CHEWED	47	CHILL	57	153
19.	/t/	TIP	24	TORE	26	TOT	46	TILL	58	154
20.	/ʃ/	SHOES	30	SHIN	33	SHOWN	39	SHIP	60	162
21.	/dʒ/	JET	27	GYP	46	JAR	50	JUDE	60	183

TERMINAL CONSONANT STIMULI

Rank	Phoneme	Individual Tokens and Numbers of Errors								Totals
1.	/n/	BUN	8	KIN	11	ROAN	14	MAIN	16	60
2.	/t/	LIGHT	11	COAT	12	KIT	13	SIT	27	63
3.	/d/	ROAD	11	RID	17	BUD	19	WED	20	67
4.	/b/	RIB	5	COB	11	DUB	27	ROBE	30	73
5.	/p/	RIP	10	ROPE	16	CUP	20	PUP	32	78
5.	/m/	SUM	11	RIM	12	CAME	26	ROAM	29	78
6.	/k/	COKE	10	BUCK	20	WICK	36	LEAK	55	121
7.	/f/	LIFE	24	ROOF	25	LEAF	35	REEF	48	132
8.	/g/	LAG	26	DUG	29	BUG	36	LEAGUE	55	149
9.	/z/	BUZZ	30	DOZE	32	SEIZE	50	LIES	62	174
10.	/v/	DOVE	30	WAVE	40	LIVE	53	LOAVE	63	186
11.	/ð/	LOATHE	18	SOOTHE	45	TEETHE	66	LITHE	84	213
12.	/s/	BUS	55	LOOSE	57	LEASE	61	LICE	62	235

APPENDIX C

STIMULUS LISTS YIELDING APPROXIMATELY EQUAL SCORES IN THE MAIN STUDY

STIMULUS LISTS YIELDING APPROXIMATELY EQUAL SCORES IN THE MAIN STUDY

Phoneme		List 1	List 2	List 3	List 4
Initial....					
1.	/p/	PILL	PAN	PAT	PAUL
2.	/t/	TILL	TOT	TORE	TIP
3.	/k/	CAT	COAL	KILL	COT
4.	/tʃ/	CHAR	CHILL	CHEWED	CHEST
5.	/b/	BAIL	BOWL	BORE	BET
6.	/d/	DEED	DO	DAB	DIE
7.	/g/	GOAL	GOT	GALE	GILL
8.	/dʒ/	JUDE	JET	GYP	JAR
9.	/m/	MORE	MAIL	MILL	MUTT
10.	/n/	KNOCK	NEAR	NEW	NAIL
11.	/f/	FAULT	FEEL	FINE	FEAR
12.	/θ/	THINK	THICK	THIGH	THOUGHT
13.	/s/	SINK	SOUGHT	SAD	SOWN
14.	/ʃ/	SHOWN	SHIP	SHIN	SHOES
15.	/h/	HOLE	HAT	HEAT	HUM
16.	/v/	VEIL	VINE	VEST	VAT
17.	/ð/	THEN	THY	THAT	THINE
18.	/z/	ZONE	ZOO	ZEAL	ZED
19.	/w/	WENT	WAIL	WORE	WIND
20.	/r/	RATE	RUES	RIP	ROT
21.	/l/	LIP	LEAP	LOT	LOSE
22.	/j/	YACHT	YOU	YALE	YEAR
Terminal...					
23.	/p/	ROPE	CUP	PUP	RIP
24.	/t/	SIT	LIGHT	KIT	COAT
25.	/k/	BUCK	COKE	WICK	LEAK
26.	/b/	RIB	ROBE	COB	DUB
27.	/d/	BUD	WED	ROAD	RID
28.	/g/	LEAGUE	BUG	DUG	LAG
29.	/m/	ROAM	SUM	RIM	CAME
30.	/n/	KIN	MAIN	BUN	ROAN
31.	/f/	LEAF	ROOF	REEF	LIFE
32.	/s/	LOOSE	BUS	LICE	LEASE
33.	/ð/	LOATHE	TEETHE	LITHE	SOOTHE
34.	/v/	LIVE	LOAVE	DOVE	WAVE
35.	/z/	LIES	DOZE	BUZZ	SEIZE

STIMULUS LISTS YIELDING APPROXIMATELY EQUAL SCORES IN THE MAIN STUDY (CONT'D.)

Phoneme	List 1	List 2	List 3	List 4
Vowel....				
36. /i/	KEEP	LEAK	TEAM	PEAT
37. /I/	BIG	KIN	SIP	BID
38. /ɛ/	NET	PECK	CHECK	PEP
39. /æ/	CAT	PAT	BAG	SAP
40. /ɝ/	HURT	BIRD	LURK	TERM
41. /ɑ/	COP	LOCK	TOM	POT
42. /ɔ/	BOUGHT	CAWED	DAWN	TALK
43. /ʌ/	PUN	BUT	DUCK	CUD
44. /ʊ/	COULD	PUT	HOOK	SOOT
45. /u/	BOOT	DUNE	LUKE	COOP
Diphthong...				
46. /aɪ/	NINE	RISE	TIDE	PIES
47. /ɔɪ/	POISE	LLOYD	BOYS	TOYED
48. /aʊ/	BOUT	LOUD	SOWS	FOUND
49. /oʊ/	ROSE	BEAUS	ROSE	TONE
50. /eɪ/	TALE	BAIT	GATE	PAYS

LIST OF REFERENCES

Beasley, W.C. National Health Survey: Hearing Study Bulletin, V. Washington, D.C.: U.S. Public Health Service, 1938.

Carhart, Raymond. "Monitored Live-Voice as a Test of Auditory Acuity," Journal of the Acoustical Society of America, XXII (1946a), 339-49.

________. "Speech Reception in Relationship to Patterns of Pure-Tone Loss," Journal of Speech Disorders, XI (1946b), 97-108.

Castle, William. The Effect of Narrow Band Filtering on the Perception of Certain English Vowels. Hague: Mouton and Co., 1964.

Chiba, T., and Kajiyama, M. The Vowel. Tokyo: Phonetic Society of Japan, 1938.

Chomsky, N., and Halle, M. "Some Controversial Questions in Phonological Theory," Journal of Linguistics, 1 (1965), 97-138.

________, and ________. The Sound Pattern of English. New York: Harper & Row, 1968.

Delattre, Pierre; Liberman, Alvin M.; Cooper, Franklin S. "Acoustic Loci and Transitional Cues for Consonants," Journal of the Acoustical Society of America, XXVII (1955), 769-73.

________; ________; ________; and Gerstman, Louis J. "An Experimental Study of the Acoustic Determinants of Vowel Color; Observations on One- and Two-Formant Vowels Synthesized from Spectrographic Patterns," Word, VIII (1952), 195-210.

Dudley, H. "The Vocoder," Bell Laboratories Record, XVIII (1939), 122-6.

Egan, J.P. "Articulation Testing Methods," Laryngoscope, LVIII (1948), 955-91.

Fairbanks, Grant. "Tests of Phoneme Differentiations: The Rhyme Test," Journal of the Acoustical Society of America, XXX (1958), 596-600.

Fletcher, Harvey. Speech and Hearing. New York: D. Van Nostrand, 1929.

________. "A Method of Calculating Hearing Loss for Speech from an Audiogram," Journal of the Acoustical Society of America, XXII (1950), 1-5.

________. Speech and Hearing in Communication. New York: D. Van Nostrand, 1953.

________, and Steinberg, J.C. "Articulation Testing Methods," Bell System Technical Journal, VIII (1929), 806-54.

Flower, Richard. Final Technical Report on Helium Speech Investigations. Report by Singer, Kearfott Division, Pleasantville, New York, to the Office of Naval Research; March, 1968 through November, 1969. Pleasantville: Singer, Kearfott Division, 1969.

French, N.R., and Steinberg, J.C. "Factors Governing the Intelligibility of Speech Sounds," Journal of the Acoustical Society of America, XIX (1947), 90-119.

Fry, N.B., and Kerridge, P.M. "Tests for the Hearing of Speech and Deaf People," Lancet, CCXXXVI (1939), 106-9.

Gay, Thomas. A Perceptual Study of American English Diphthongs. Supplement to Status Report on Speech Research, New York: Haskins Laboratories, 1967.

________. "Effect of Speaking Rate on Diphthong Formant Movements," Journal of the Acoustical Society of America, XLIV (1968), 1570-3.

Griffiths, John. "Rhyming Minimal Contrasts: A Simplified Diagnostic Articulation Test," Journal of the Acoustical Society of America, XLII (1967), 236-41.

Haggard, M., and Mattingly, I.G. "A Simple Program for Synthesizing British English," IEEE Transactions on Audio and Electroacoustics, 1968.

Harris, K.S. "Cues for the Discrimination of American English Fricatives in Spoken Syllables," Language and Speech, I (1958), 1-7.

Heinz, John, and Stevens, Kenneth. "On the Properties of Voiceless Fricative Consonants," Journal of the Acoustical Society of America, XXXIII (1961), 589-96.

Hirsh, Ira. The Measurement of Hearing. New York: McGraw-Hill, 1952.

________; Davis, H.; Silverman, R.; Reynolds, E.G.; Eldert E.; and Benson, R. "Development of Materials for Speech Audiometry," Journal of Speech and Hearing Disorders, XVI (1952), 321-38.

House, Arthur S.; Williams, Carl E.; Hecker, Michael H.L.; and Kryter, K.D. "Articulation-Testing Methods: Consonantal Differentiation with a Closed-Response Set," Journal of the Acoustical Society of America, XXXVII (1965), 158-66.

Hutton, Charles, and Weaver, J. "PB Intelligibility and Word Familiarity," Laryngoscope, LXIX (1959), 1443-50.

Jakobson, R.; Fant, G.; and Halle, M. Preliminaries to Speech Analysis. Cambridge: MIT Press, 1963.

Klatt, Dennis, H. "Structure of Confusions in Short-Term Memory between English Consonants," Journal of the Acoustical Society of America, XLIV (1968), 401-7.

Kryter, K.; Williams, C.; and Green, D. "Auditory Acuity and the Perception of Speech," Journal of the Acoustical Society of America, XXXIV (1962), 1217-23.

Lehiste, I., and Peterson, G. "Linguistic Considerations in the Study of Speech Intelligibility," Journal of the Acoustical Society of America, XXXVI (1959), 280-6.

Lehmann, W.P., and Heffner R.M.S. "Notes on the Length of Vowels (III)," American Speech, XV (1940), 377-80.

Liberman, A.M.; Delattre, P.; and Cooper, F.S. "The Role of Selected Stimulus Variables in the Perception of the Unvoiced Stop Consonants," American Journal of Psychology, LXVI (1952), 497-516.

________; ________; ________; and Gerstman, Louis J. "The Role of Consonant-Vowel Transitions in the Perception of the Stop and Nasal Consonants," Psychological Monographs, LXVIII (1954), 1-13.

________; ________; Gerstman, L.J.; and Cooper, F.S. "Tempo of Frequency Change as a Cue for Distinguishing Classes of Speech Sounds," Journal of Experimental Psychology, LII (1956), 127-37.

Licklider, J.C.R., and Miller, George A. "The Perception of Speech," Handbook of Experimental Psychology, edited by S.S. Stevens. New York: John Wiley & Sons, 1951.

Lisker, Leigh. "Minimal Cues for Separating /w,r,l,y/ in Intervocalic Position," Word, XIII (1957), 256-67.

________, and Abramson, A. "A Cross-Language Study of Voicing in Initial Stops: Acoustical Measurements," Word, XX (1964), 384.

________, and ________. "Some Effects of Context on Voice Onset Time in Eglish Stops," Language and Speech, X (1967), 1-28.

Locke, W.N. and Heffner, R.M.S. "Notes on the Length of Vowels (II)," American Speech, XV (1940), 74-9

Martin, Frederick Noel. An Investigation of Transcranial Stimulation of Supra-Liminal Speech Stimuli in Mixed Hearing Loss. Unpublished Ph.D Dissertation, City University of New York, 1968.

________; Hawkins, R.; and Bailey, H. "The Nonessentiality of the Carrier Phrase in PB Word Testing," Journal of Auditory Research, II (1962), 319-22.

Mattingly, Ignatius G. Synthesis by Rule of General American English. Supplement to Status Report on Speech Research, New York: Haskins Laboratories, 1968.

Miller, G.A. "The Perception of Speech," For Roman Jakobson, edited by M. Halle. Hague: Mouton & Co., 1956.

________, and Nicely, Patricia E. "An Analysis of Perceptual Confusions among Some English Consonants," Journal of the Acoustical Society of America, XXVII (1955), 338-52.

Mullins, Cecil, and Bangs, J.L. "Relationships between Speech Discrimination and Other Audiometric Data," Acta Otolarygologica, XLVII (1957), 149-57.

O'Connor, J.D.; Gerstman, L.J.; Liberman, A.M.; Delattre, P.C.; and Cooper, F.S. "Acoustic Cues for the Perception of Initial /w,j,r,l/ in English,"Word, XIII (1957), 24-43

O'Neill, John, and Oyer, Herbert. Applied Audiometry. New York: Dodd, Mead & Co., 1966.

Owens, Elmer; Talbott, Carolyn; and Schubert, Earl. "Vowel Discrimination of Hearing-Impaired Listeners," Journal of Speech and Hearing Research, XI (1968) 648-55.

Oyer, Herbert, and Doudna, Mark. "Structural Analysis of Word Responses Made by Hard-of-Hearing Subjects on Discrimination Tests," Archives of Otolaryngology, LXX (1959), 357-64.

Peterson, G.E. "The Information Bearing Elements of Speech," Journal of the Acoustical Society of America, XXIV (1952), 629-37.

________, and Barney, Harold. "Control Methods Used in a Study of the Vowels," Journal of the Acoustical Society of America, XXIV (1952) 175-84.

________, and Lehiste, I. "Duration of Syllable Nuclei in English," Journal of the Acoustical Society of America, XXXII (1960), 693-703.

________, and ________. "Revised CNC Lists for Auditory Tests," Journal of Speech and Hearing Disorders, XXVII (1962). 61-70

Pickett, J.M. "Perception of Vowels Heard in Noises of Various Spectra," Journal of the Acoustical Society of America, XXIX (1957), 613-20.

________; Martin, E.S.; Johnson, D.; Smith, S.B.; Daniel, Z.; Willis, D.; and Otis, W. "On Patterns of Speech Feature Reception by Deaf Listeners," Speech Communication Ability and Profound Deafness, edited by G. Fant. Washington, D.C.: The Alexander Graham Bell Assn. for the Deaf, 1972, 119-33.

Plummer, Robert. "High Frequency Deafness and Discrimination of High-Frequency Consonants," Journal of Speech Disorders, VII (1943) 373-81.

Pollack, Irwin. "Effects of High-Pass and Low-Pass Filtering on the Intelligibility of Speech in Noise," Journal of the Acoustical Society of America, XX (1948), 259-66.

________; Rubenstein, Herbert; and Decker, Louis. "Intelligibility of Known and Unknown Message Sets," Journal of the Acoustical Society of America, XXXI (1959), 273-9.

Potter, R.K., and Steinberg, J.C. "Toward the Specification of Speech," Journal of the Acoustical Society of America, XXII (1950), 807-820.

Rhodes, Robert C. "Discrimination of Filtered CNC Lists by Normals and Hypacusics," Journal of Auditory Research, VI (1966) 129-33.

Rosen, Jack. Phoneme Identification in Sensorineural Deafness. Unpublished Ph.D. Dissertation, Stanford University, 1962.

Savin, Harris B. "Word-Frequency Effect and Errors in the Perception of Speech," *Journal of the Acoustical Society of America*, XXXV (1963), 200-6.

Schultz, M.C. "Suggested Improvements in Speech Discrimination Testing," *Journal of Auditory Research*, IV (1964), 1-14.

Siegenthaler, Bruce. "A Study of the Relationship between Measured Hearing Loss and the Intelligibility of Selected Words," *Journal of Speech and Hearing Disorders*, XIV (1949), 111-8.

Speaks, Charles. "Intelligibility of Filtered Synthetic Sentences," *Journal of Speech and Hearing Research*, X (1967),289-98.

Steinberg, J.C.; Montgomery, H.C.; and Gardner, M.D. "Results of the World's Fair Hearing Tests," *Journal of the Acoustical Society of America*, XII (1940), 290-301.

Studebaker, Gerald A. *The Audiological Application of Tests of Phonemic Differentiation*. Project Report, University of Oklahoma, 1968.

Thorndike, E.L., and Lorge, I. *The Teacher's Word Book of 30,000 Words*. New York: Teachers College, Columbia University, 1944.

Utley, Jean. "The Relation between Speech Sound Discrimination and a Percent of Hearing Loss," *Journal of Speech Disorders*, IX (1944), 106-13.

Webster, J.C., and Klumpp, R.G. "Articulation Index and Average Curve-Fitting Methods of Predicting Speech Interference," *Journal of the Acoustical Society of America*, XXXV (1963), 1339-44.

Wickelgren, Wayne. "Distinctive Features and Errors in Short-Term Memory for English Vowels," *Journal of the Acoustical Society of America*, XXXIX (1965a), 583-8

________. "Distinctive Features and Errors in Short-Term Memory for English Consonants," *Journal of the Acoustical Society of America*, XXXIX (1965b) 388-98.

Young, Martin, and Gibbons, Edward. "Speech Discrimination Scores and Threshold Measurements in a Non-Normal Hearing Population," *Journal of Auditory Research*, II (1962), 21-33.