A PHONETIC ANALYSIS OF TIME-COMPRESSED CNC MONOSYLLABLES

Thesis for the Degree of M. A.
MICHIGAN STATE UNIVERSITY
D. CREIG DUNCKEL
1972

THESIS

Accepted by the faculty of the Department of Audiology and Speech Sciences, College of Communication Arts, Michigan State University, in partial fulfillment of the requirements for the degree of Master of Arts.

Thesis Committee:

Daniel S. Beasley, Ph. D. — Director

William F. Rintelmann, Ph. D.

May Chin

ABSTRACT

A PHONETIC ANALYSIS OF TIME-COMPRESSED CNC MONOSYLLABLES

By

D. Creig Dunckel

The purpose of this study was to investigate the effects the time-compression procedure had on the intelligibility of CNC monosyllables.

The experimental stimuli utilized was list 1, form B, of the Northwestern University Auditory Test Number 6. The words in this list were time-compressed by 30% through 70%, in 10% steps, in addition to a 0% control condition. Compression was accomplished with the Zemlin modification of the Fairbanks Time-Compressor. The data for this study were gathered by Beasley *et al*. (1972).

Confusion matrics were made for the errors associated with each degree of time-compression (0%, 30%, 40%, 50%, 60%, and 70%) and each sensation level of presentation (8 dB, 16 dB, 24 dB, and 32 dB). Thus a total of 24 confusion matrices were recorded. These matrices were then condensed over all conditions of

time-compression levels. They were further classified by sensation levels, with 8 and 16 dB combined and labelled as low sensation levels, and 24 and 32 dB combined and labelled as high sensation levels. In each matrix the following phonemes were considered: / p, b, t, d, k, g, f, v, θ, s, z, ʃ, m, n, tʃ, dʒ, r, j, h, l, hw, ŋ, l. The consonants were further classified according to the linguistic features of voicing, nasality, duration, and place of articulation.

Results indicated that a consistency in phonemic errors does exist and may be due to a change in the speech signal by various degrees of time-compression. Further, these phonemic confusions differ depending upon the placement of the phoneme in the word i.e. initial vs final position. Different substitution patterns were exhibited for the same phoneme when in the initial position of the word as opposed to the final position.

The sensation level of presentation also affected the type of phonemic errors. This was demonstrated by a / b / for / v / substitution at low sensation levels and / m / for / v / substitution at the high sensation levels.

This study further revealed that the distinctive features of the phonemes were also affected by the time-compression procedure. Those phonemes with the greatest duration tended to be least affected by the various time-compression ratios.

A PHONETIC ANALYSIS OF TIME-COMPRESSED CNC MONOSYLLABLES

By

D. Creig Dunckel

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

MASTER OF ARTS

Department of Audiology and Speech Sciences

1972

ACKNOWLEDGMENTS

I wish to extend my gratitude to Dr. Daniel S. Beasley, my thesis advisor, and to Dr. William F. Rintelmann and Professor May Chin, the members of my committee, for their assistance in the preparation of this thesis.

I would also like to extend my appreciation to Miss Shelley Schwimmer for the collection of the data used in this thesis.

Last, but not least, I wish to acknowledge my wife, Priscilla, for her faith and endurance in helping me achieve my educational goals.

TABLE OF CONTENTS

Chapter Page

LIST OF TABLES

Table | Page

Table		Page

CHAPTER I

INTRODUCTION

Audiological evaluation techniques have concerned themselves primarily with detection of lesions in the peripheral auditory system and in the eighth cranial nerve. Recently, however, there has been an increasing interest in detection of lesions in the brain-stem and auditory cortex. The structure and sensitivity that is required of a test to measure lesions in the higher auditory centers, however, are lacking in present conventional tests.

Individuals with disorders of the higher auditory pathways usually exhibit response behavior which appears essentially normal (Willeford, 1969; Katz, 1969). The function of higher auditory centers is to organize simultaneous or successive elements of the acoustic speech signal into a definite pattern (Bocca and Calearo, 1963). It is recognized that the central auditory pathways provide a sufficient degree of intrinsic redundancy, even when damaged, to allow simple psycho-acoustic elements to satisfactorily reach the cortical integrative and interpretive centers (Bocca and Calearo, 1963). Attention has therefore become centered on tests involving verbal stimuli which have been altered to render comprehension more difficult by adequately taxing this intrinsic redundancy by decreasing extrinsic redundancy.

In an effort to produce verbal stimuli which renders comprehension more difficult, stimulus distortion, such as described by Fairbanks, Everitt, and Jaeger (1954), has been used. A study by Beasley, Schwimmer, and Rintelmann (1972) using ninety-six right-handed normal hearing young adults, sought to isolate the effects of varying degrees of time-compression on monosyllabic word intelligibility and to examine them with respect to intensity increases in the speech stimuli. Their study provided the groundwork for further studies, including investigations of auditory pathologies. The results of this study showed that increases in time-compressed CNC monosyllables to 60% resulted in gradual decreases in intelligibility. At 70% time-compression a sharp decline in intelligibility occurred. Also, intelligibility increased as sensation level increased for 8 dB SL to 32 dB SL. The results of this study tend to show that time-compressed speech does render comprehension more difficult. Beasley *et al*. suggested that such stimuli may assist in diagnosing higher auditory pathway lesions.

Fournier (1954) stressed the importance of the time factor in speech discrimination and related the increase in time required by the cortex for identification of a message to the difficulties in speech perception experienced by the elderly. Bordley and Haskins (1955) demonstrated that an increase in rate of presentation of the word "stimulus" resulted in more difficulty in comprehension. Calearo and Lazzaroni (1957) did a study of precise discrimination scores of normal and presbycusic subjects, varying the factors of intensity and syllabic rate of "short, significant sentences" as stimulus material. They demonstrated that in normal subjects, an increase in syllabic rate is almost completely neutralized by a simultaneous increase in intensity. When the accelerated sentence material was presented to the group of presbycusic subjects, however, threshold shift was increased as much as 30 dB for the intermediate speed, while a threshold of speech perception could not be obtained at any intensity level for the highest speed. Furthermore, it was reported that when the accelerated sentences were presented to a group of subjects with temporal lobe lesions, these subjects yielded poor articulation curves when the speeded message was sent to the ear contralateral to the lesion.

de Quiros (1964) used accelerated speech material to test 20 normal subjects, 15 subjects with peripheral hearing losses, seven presbycusics, and several groups of adults and children with central disorders. The speech material used in his study were "abstract" sentences of approximately ten words for adult subjects, and "concrete" sentences of similar length for use with children. The subjects' articulation scores were considered in relation to the shape of the articulation curve, speech detection threshold, speech reception threshold, and maximum articulation score. Results obtained with the groups of subjects with CNS disorders indicated that in differential diagnosis, accelerated speech testing provided additional information, which, when correlated with other findings, may aid in pinpointing sites of brain lesions, especially within the temporal lobe. However, because of the small number of subjects considered with CNS disorders, the consistency of responses from subject to subject within each category and the differential responses of subjects in each of the categories cannot be considered definitive on the basis of this study.

Controlled Time-Compression Procedures

Fairbanks, Everitt, and Jaeger (1954) produced a controlled electromechanical procedure for compressing speech stimuli. The procedure involved passing a tape over the curved surface of a cylinder and wrapping it around the cylinder enough to make contact with one-quarter of its circumference. This was accomplished by four tape reproducing heads equally spaced around the circumference of the cylinder. When this cylinder is stationary, and the tape is moving at the same speed at which it moved during recording, it makes contact with one of the reproducing heads and the signal is reproduced as recorded. When an adjustment is made for a certain amount of compression, the speed of the tape increases and the cylinder begins to rotate in the direction of the tape motion. Under conditions of time-compression, each of the four heads makes, and then loses, contact with the tape loop. Each head reproduces, as recorded, the material on that portion of the tape with which it makes contact. When the cylinder is so positioned that one head is just losing contact with the tape while the preceding head is just making contact with the tape, the segment of the tape that is wrapped around the cylinder between these two heads does not make contact with a reproducing head, and is therefore not reproduced. The information not reproduced is referred to as the interval of discard, and was found to be maximally effective when it was 15 to 20 m/sec in length (Fairbanks and Kodman, 1957). The amount of speech compression is dependent upon the number of discard intervals (non-reproduced segments) per unit of time.

The sampling interval (Ia) is composed of the discard interval (Id) and the recorded interval (Ir), such that

$$Ia = Id + Ir$$

The sampling frequency (Fs), i.e., the rate at which the input signal is sampled, is

$$Fs = 1/Ia$$

The compression ratio (Rc) is then defined as

$$Rc = Id/Ia = IdFs$$

Both Rc and Id are independently manipulated since, on the Fairbanks apparatus, the tape and cylinder speed are independently variable. This allows for variation of the temporal value of the discarded portions of the message.

Using this electromechanical apparatus, Fairbanks, Guttman, and Miron (1957) presented a pair of independent message-test units to normal hearing subjects. Each consisted of an extended exposition of technical information and a corresponding test of factual comprehension. The messages were read and recorded approximately 141 wpm and time-compressed electro-mechanically by various amounts. Independent groups of subjects were assigned to five experimental conditions which represented a series of compressions from 0% to 70%, and to a sixth test condition in which no message was presented. Listener aptitude was controlled by forming subgroups of approximately "equal" aptitude for each condition at four different levels. The effect of the message-test difficulty was assessed by sub-scoring resulte according to five message effectiveness levels, based upon differences in responses to test items in the 0% compression condition and the test-only condition. The curve of comprehension as a function of message time was found to be sigmoid. Response scores were approximately 50% of the maximum when the message was compressed by 60%, and slightly less than 90% of the maximum when the message was compressed by 50%. It was concluded that the interaction of time-compression and message effectiveness significantly affected comprehension of factual material.

To isolate the effect of reduced stimulus duration on intelligibility Fairbanks and Kodman (1957) presented Egan's phonetically balanced (PB-50) word lists to highly trained listeners. The words were presented at a constant intensity at varying ratios of time-compression. The results of their study demonstrated that time-compression, up to 20% of the original signal duration, had no significant influences on intelligibility, when the stimulus words were presented at a comfortable listening level.

It has been demonstrated that the PB-50 word lists reflect low reliability and a wide range of word difficulty among the test items (Eldert and Davis, 1951). Although the W-22 word lists were devised (Hirsh et al., 1952) in an attempt to overcome the reliability problems in discrimination testing, it was not until the development of the Northwestern University Auditory Test No. 6 (NU-6) that the latter difficulty was adequately solved (Tillman and Carhart, 1966).

The development of this test was based upon studies performed with an earlier test, the Northwestern Auditory Test No. 4 (NU-4) developed by Tillman, Carhart, and Wilbur (1963). This test was used extensively in the Auditory Research Laboratories at Northwestern University for a two-year period. It proved to be a valuable tool in the measurement of speech discrimination. The NU-4 consisted of six randomizations of two 50-word lists composed of phonemically balanced monosyllables of the consonant-nucleus-consonant (CNC) variety, selected from a pool of such words compiled by Lehiste and Peterson (1959). In addition, the 50 words in each list contained a proportional distribution of phonemes in the Thorndike and Lorge (1952) word list. It was found that even with six equivalent forms of each list, the investigation of a large number of listening conditions could not be accomplished without several repetitions of the various forms and lists. Because of this limitation the NU-4 was revised and expanded into the NU-6, which consists of four randomizations of four phonemically balanced word lists, each composed of 50 monosyllables. From the studies of Tillman and Carhart (1966), it was found that NU-6 compares favorably with the NU-4 in interlist equivalence and test-retest reliability. Further as with the NU-4, subjects with conductive hearing losses yielded articulation functions closely duplicating normal subjects, whereas, subjects with sensori-neural impairment yielded more gradually rising articulation functions and lower mean maximum articulation scores. The NU-6, therefore, satisfies the two basic requirements of a diagnostically useful test of auditory discrimination: (a) a substantial segment of the articulation function, depicted graphically as an articulation curve, is linear, so that the value of its slope may be precisely measured, and (b) the slope of the articulation function has the potential to diagnostically differentiate among various auditory pathologies.

Phonetic Considerations in Perceptual Processing

The acoustic message is correlated to sounds which can be characterized by their amplitude and frequency in the case of a pure tone, or by the amplitude and spectrum in the case of a complex sound. The physical quality of frequency corresponds to the sensation of pitch, the amplitude to the sensation of loudness, and the spectrum to the sensation of timbre or quality.

While a sound must be perceived to be identified, it is by no means certain that once a sound and/or word is perceived, it will be identified. Thus the object of vocal audiometry is to study the intelligibility of speech.

As Malmberg (1970) points out, the clarity of the speech message is very important for the identification of the message. The intensity has a very definite effect on the clarity. Experiments carried out with a message spoken at normal intensity and then played back with various attenuations show that discrimination is optimum between 60 and 70 dB SPL, and becomes more difficult above 80 dB SPL (Licklider and Miller, 1938). Licklider and Miller (1938) also pointed out that information pertaining to the duration of the message and the distortion involved in the production of the message are essential in determining the intelligibility of the message. "Speeded-up" speech thus reduces the possibilities of identification even if the frequencies involved are not essentially modified.

A form of measuring distorted speech is via filtering. The first studies of clarity as a function of distortion in the form of filtering were carried out by Fletcher (1929). He showed that in order to retain 50% of the phonetic clarity the cut-off frequency of a filter should not exceed 1200 Hz for a high pass filter, or be less than 1700 Hz for a low pass filter.

Lafon (1961) showed that the replacement of one speech sound by another in a discrimination task is not a random process, but depends on the significant acoustic feature in question. Contoids are more often incorrectly heard than vocoids. As regards the plosives, / k / is usually mistaken for a / t /, or sometimes for a / p /. The / p /, whose explosion is less accentuated may be mistaken for the corresponding voiced contoid, / b /. The / r / may be confused for all voiced contoids without any particular preference except perhaps for / l /; but practically never replaces another speech sound. The / l / and the / m / are most often replaced by / n /, and / j / is mistaken for / l / and / n /. Further, Lafon (1961) found that the labio-dental / f / is most often heard incorrectly and the sibilants are most often mistaken for one another, typically in the form of voiced for voiced and unvoiced for unvoiced confusions.

Duration of the Speech Signal. The duration of the speech signal helps to distinguish certain phonemes from other phonemes. For example, the plosive consonants / p, t, k, b, d, g, / are characterized by a burst.

Further, the voiced plosives / b, d, g / consume more time than their voiceless counterparts / p, t, k /. Fricatives and affricatives / f, θ, s, ʃ, tʃ, v, ð , z, ʒ , and dʒ / are distinguished by friction and relatively greater duration than plosives. All durational phenomenon can be demonstrated spectrographically (Jakobson et al., 1952).

The durational aspect of the consonants contributes to distinguishing plosives from fricatives. Voiceless plosives require a minimum of temporal clues for recognition, followed by the voiced plosives, slit fricatives, and the grooved fricatives, respectively. Black and Singh (1970) reported that the errors made by a panel of listeners, when increasing amounts of the initial portions of consonants were removed, consisted of / f / and / v / being confused as / b /, / s / as / t /, / z / as / d /, / ʃ / as / t /, and / ʒ / as / d /.

In contrast, Tiffany (1953) compared the intelligibility of sections of vowels of 0.08, 0.2, 0.5 and 8.0 sec, and concluded that added duration, beyond a "natural duration" in speech does not contribute to recognition. This result was in agreement with Siegenthaler's (1950) findings that ten "long-sustained" vowels were only 50% recognizable.

Miller and Licklider (1940) using a method of deleting the alternate sections of recorded speech without affecting the over-all duration of the passage demonstrated that much of the continuous time consumed by words can be deleted without seriously impairing the perception of the words. They varied both the frequency and the duration of the interruption of the signal and used recognition of monosyllables as a criterion measure. With one interruption per second and with the interruption lasting 0.5 sec, the words were 40% intelligible, and when the interruption was 0.25 sec they were 80% intelligible. With ten interruptions per second and only 25% of the word remaining, the reception score was 60%, thus showing that as more of the speech signal was removed, intelligibility decreased. Further, Garvey and Henneman (1952) found that more than 60% of the speech pattern had to be removed before intelligibility decreased below 80%. It was suggested by Daniloff, Shriner, and Zemlin (1968) that at high compression ratios, a normal listener may not have enough time or information to perceptually process incoming verbal stimuli correctly thus lowering the intelligibility of the signal. Neither of these studies attempted to make an analysis of the errors made by the listeners.

In essence, the duration of the signal, whether it is lengthened or shortened, plays a role in the listener's ability to integrate the signal in the higher auditory pathways. Many authors (Fletcher and Steinberg, 1929; Black, 1952; Stevens, 1946; and Miller, Heise, and Lichten, 1951) have found that accurate perception of a spoken word also depends to a considerable extent upon the frame of expectations within which a word occurs and the length of the signal. Thus, nonsense words may be less intelligible than meaningful units and monosyllables may be less intelligible than polysyllabic units and/or sentences, up to a point (Beasley and Shriner, 1972).

Intensity of the Speech Signal. The intensity or the level of presentation of the signal plays an important role in the listener's ability to perceive and interpret the signal. Beasley, Schwimmer, and Rintelmann (1972) presented time-compressed stimuli at sensation levels of 8, 16, 24, and 32 dB and demonstrated that as the sensation level of presentation increased the intelligibility of the stimuli increased. Further, the largest increase in intelligibility due to an increase in the intensity of the signal occurred between 8 and 16 dB sensation level.

Miller and Nicely (1954) used six different signal to noise ratios ranging from -12 to +12 dB in 6 dB steps. They demonstrated that as the S/N ratio increased, the intelligibility of the stimuli increased.

Stevens (1938) investigated psychophysical data for pure tones as related to intensity, but unfortunately there is practically no psychophysical data relating the loudness of speech sounds to their intensity except for some general observations made by Fletcher (1929). He found that sounds with a large number of frequency components increase more rapidly in loudness with a rise in intensity than do sounds with fewer components. Further, he noted that sounds with most of their energy in the low frequency region increase more rapidly in loudness than those with high frequency energy concentrations. Thus the loudness of any complex sound will depend upon its frequency components with the vowels being perceived louder than consonants because of the point of their energy concentration. That is, words that have sounds with most of their energy in the low frequency region may be perceived as louder, thus increasing their intelligibility, than words whose sounds have energy in the high frequencies.

Linguistic Features of Consonants. There are many ways of classifying phonemes according to features and the articulation process used to generate these sounds. These features help the listener to distinguish one phoneme from another and thus enhance the intelligibility of the phoneme. According to one theory, in order to decode the message, the receiver extracts the distinctive features from the acoustical data (Jakobson and Halle, 1970) in order to make perceptual decisions.

Distinctive features can be broken into two classes: the prosodic features, including force, quality, and tone, and the inherent features which are divided into three subclasses: sonority, protensity, and tonality. Sonority, protensity, and tonality correspond to the prosodic features of force, quality, and tone (Jakobson and Halle, 1970). The sonority features are classified in the following manner: (1) vocalic/non-vocalic, (2) consonantal/non-consonantal, (3) nasal/oral, (4) compact/diffuse, (5) abrupt/continuant, (6) strident/non-strident (mellow), (7) checked/unchecked, (8) voice/voiceless. The protensity feature is tense/lax and the tonality features are grave/acute, flat/non-flat, and sharp/non-sharp (Jakobson and Halle, 1970).

Miller and Nicely (1955) in an experiment using English consonants transmitted with frequency distortion and random masking, confirmed that the perception of each of these features is relatively independent of the perception of the others. This study involved 16 consonants followed by the vowel / a / as in "father." These phonemes were recorded under six conditions of signal-to-noise ratios beginning with -12 to +12 dB, in 6 dB steps. Different frequency responses were measured at the +12 dB S/N condition.

In this study the 16 consonants were classified according to the following linguistic features:

(1) voiced/voiceless, including / b, d, g, v, ð, z, ʒ, m, n, / as voiced phonemes and / p, t, k, f, θ, s, ʃ, / as unvoiced phonemes;

(2) nasality, including / m / and / n /;

(3) affriction, classified according to articulatory production, including / f, θ, s, ʃ, v, ð, z, ʒ, /;

(4) duration, whereby / s, ʃ, z, ʒ, / were separated from the other 12 consonants because of the former's added duration and

(5) place of articulation, according to tongue placement in front, middle, or back of the oral cavity, with / p, b, f, v, m, / considered as front, / t, d, θ, s, ð, z, n, / as middle, and / k, g, ʃ, ʒ, / as back consonants.

The Miller and Nicely (1955) study showed that these five linguistic features could be effectively used for grouping consonants, and that their distinctive features do aid in the identification of one phoneme from another under varying signal-to-noise conditions. The results indicated that voicing and nasality are much less affected by a random masking noise than the other features. The results for affrication and duration were very similar and superior to place of articulation for auditory discrimination, but inferior to nasality and voicing. Place of articulation was difficult to discriminate at ratios less than +6 dB whereas nasals and voicing could be discriminated at signal-to-noise ratios of -12 dB.

Summary and Statement of the Problem

In summary, a review of the literature suggests that time-compressed speech stimuli may have value as a diagnostic tool for identifying lesions of the higher auditory pathways (Bocca and Calearo, 1963; Calearo and Lazzaroni, 1957; deQuiros, 1964). However, several studies reveal that when distortion of the speech signal occurs by either altering the duration of the signal and/or the intensity of the signal, certain phonemes may be confused with other phonemes (Lafon, 1961; Jakobson et al., 1952; Black and Singh, 1970).

The purpose of this study was to analyze those errors in the form of perceptual phonemic confusions in response to stimuli time-compressed 0%, 30%, 40%, 50%, 60%, and 70%. Further, the effect of the linguistic features of voicing, nasality, duration, and place of articulation determining the intelligibility of the speech signal was investigated. In addition, perceptual confusions associated with low and high sensation levels of presentation and their effect upon phonemic confusions and signal intelligibility was studied.

Specifically, the following questions were investigated relative to the responses of normal hearing, young adult subjects to a standardized, monosyllabic word list:

1. Is there a consistency in phonemic errors due to a change in the speech signal by various degrees of time-compression (30% to 70% in 10% steps)?

2. Do these errors differ when the phoneme is in the initial or final position of the word?

3. Do the type of phonemic errors differ as a function of sensation level?

4. How are the various distinctive features affected as a result of time-compression?

CHAPTER II

EXPERIMENTAL PROCEDURES

The data for this study were obtained from a study by Beasley, Schwimmer, and Rintelmann (1972).

Subjects

The subjects were 96 normal hearing right-handed young adults selected from a university population. These subjects were randomly assigned to six groups of sixteen each. Each subject was required to pass a sweep frequency screening test presented at a Hearing Level of 22 dB (re: ISO, 1964 Standard) at octave intervals ranging from 125 Hz to 8,000 Hz to insure normative status of hearing bilaterally. Also, a live-voice presentation of the CID W-1 Word List was administered unilaterally to obtain the Speech Reception Threshold (SRT) for the designated test ear.

Stimulus Generation

The experimental stimuli used in this study were the four lists of From B of the NU-6 (Tillman and Carhart, 1966). The four-word lists were recorded at normal conversational speech and effort level by a trained white male talker who spoke General American English under controlled recording procedures (Rintelmann and Jetty, 1968).

An Ampex Model 601 tape deck (frequency response 50-12,000 Hz ±2 dB) and an Ampex Model 600-2 tape deck (frequency response 50-13,000 Hz, ±2 dB) were used to make copies of each of the four recorded lists. The copies were then temporarily processed using the Fairbanks electromechanical time-compression apparatus (Fairbanks, Everitt, and Jaeger, 1954), as modified by Zemlin (1971). Each list was time-compressed by 30%, 40%, 50%, 60%, and 70%, and was also passed through the time-compression mechanism under 0% in order to control for possible fidelity distortion when using the tapes. In all, there were six time-compressed recordings for each of the four lists, resulting in 24 experimental tape recordings.

The experimental tapes were copied using an Ampex Model 601 tape recorder and an Ampex AG 500-2 (frequency response 50-13,000 Hz, ±2 dB) monitored by an Ampex AA 620 power amplifier.

Presentation Procedures

The 96 subjects were divided into six groups, corresponding to the six different percentages of time-compression under study. Each subject within a single group was presented with the four lists of Form B of the NU-6. Each list was presented at one of four sensation levels: 8 dB, 16 dB, 24 dB, and 32 dB. The order of presentation of the four sensation levels was rotated within each group. In this manner, each test list was presented a total of four times for each time-compression condition at each sensation level, and the sensation levels were counterbalanced to avoid possible order effects of sensation level presentation.

A prefabricated double-walled test chamber (IAC 1200 series) was used in testing each subject individually. There was no interference from ambient noise in the test room since it was sufficiently low (45 dB on the C-scale of a Bruel and Kjar sound level meter) so as not to interfere at even the lowest sensation level.

Analysis

Each subject's response was recorded on an answer sheet which the experimenter hand-scored. The data were then converted to percentage correct scores and were plotted as articulation curves. The slopes of these articulation functions were calculated for each condition of time-compression. In addition, graphic data were computed for sensation level and respective interactions studied.

Procedures for the Present Investigations

Phonemic confusion matrics were made using list 1, Form B, of the N.U. Auditory Test No. 6. This list was used because of its difficulty when compared with the other lists (2, 3, and 4) of Form B (Jetty and Rintelmann, 1968). Further, it was felt that list 1, because of its difficulty, would be more representative of the errors made by the subjects. Confusion matrics were made for the errors associated with each degree of time-compression (0%, 30%, 40%, 50%, 60%, and 70%) and each sensation level of presentation (8 dB, 16 dB, 24 dB, and 32 dB). Thus a total of 24 confusion matrices were recorded. These matrices were then condensed over all conditions of time-compression levels. They were further classified by sensation levels, with 8 and 16 dB combined and labelled as low sensation levels, and 24 and 32 dB combined and labelled as high sensation levels. These matrices were further subdivided according to 20 initial and 19 final positions of the phonemes under study. In each matrix the following phonemes were considered: / p, b, t, d, k, g, f, v, θ, s, z, ʃ, m, n, tʃ, dz, r, j, h, l, hw, ŋ , /. If the subject made no response, this was recorded in the data.

The consonants were classified according to the linguistic features of voicing, nasality, duration, and place of articulation, in the manner described by Miller and Nicely (1955). / b, d, g, v, z, dʒ, m, n, ŋ, r, l, hw, j, / were classified as voiced, and / p, t, k, f, θ, s, ʃ, tʃ, h, / as voiceless. / m, n, and ŋ / were classified under nasality. Consonants classified as longer in duration were / s, ʃ, z, dʒ, /. Consonants were also classified as either front, middle, or back under place of articulation. The front consonants were / p, b, f, v, m, /; the middle consonants were / t, d, θ, s, z, n, ŋ, j, tʃ, /; and the back consonants were / k, g, ʃ, dʒ, r, l, h, hw, /. The phonemic confusion matrices were computed according to the above classification system.

CHAPTER III

RESULTS

The results of this study show certain consistencies associated with confusions at varying time-compression levels. It further reveals these confusions may differ depending upon the position of the phoneme in the word, i.e., initial vs. final. Further, the results show that the lower sensation levels resulted in more errors. The results further reveal that phonemes with certain distinctive features were more intelligible than others.

Tables 1 through 8 show the errors made at the various time-compression levels (0% through 70%). They further depict these errors by phoneme placement in the word (initial placement, final placement, and initial plus final placement). Also recorded in the tables are the percentage of time the phoneme was perceived correctly and the percentage of time it was perceived as another phoneme where applicable.

Low Sensation Level

Tables 1 through 3 reveal errors in the form of phonemic confusion matrices for the stimulus items

TABLE 1.--Phonemic confusions for phonemes in the initial position over all conditions of time-compression at low sensation levels.

	p	b	t	d	k	g	f	v	θ	s	z	S	m	n	tS	dz	r	j	h	l	hw	NR	S %	C %
p	103	5	4	1	16	1	6												1			4	k 11	72
b	2	104		5	3	10	2			1			1		1	1	1		1	2		10		72
t	3		96	5	18	1	2			1			1			1		3	4			9	k 13	67
d		6		113	2	3	1						2	1			3		1	1		11		78
k	3	2	10	10	93	6	1		1				2	1			4		1	1		9		65
g		1		20	2	65			1				1			1	2	1				2	d 21	68
f		39				1	52			1									1			2	b 41	54
v		20	1	1	1			12					6	1	1				1	2		2	b 42	25
θ		5			4	2	1		28	1								1	3			3		58
s		2	1	6						120		2					2		2	1		8		83
S				1								93					1					1		97
m		12		2	1	4	1		1				104	4			1		6	5		3		72
n	2	1		4									2	78			1			1		7		82
tS		2	3												86	4						1		90
dz		1		4	1	2									1	84			1			2		88
r		1		2			1										158			23			l 12	82
j				1		3							1					41				2		85
h		1			1	19	1												117			5	g 13	81
l		8		2						2				1		1	3		3	162		10		84
hw				2		3							1				11				116	9		81

NR = No response; S = Substitution; C = Correct

presented at the low sensation levels (8 and 16 dB) averaged over all conditions of time-compression.

Comparison of Table 1 (initial errors) and Table 2 (final errors) shows that overall, there were more final position phonemic confusions than initial position confusions. Further, reference to Tables 1A through 12A in Appendix A reveals that the major increase for initial position errors did not occur until 70% time-compression, whereas for final position errors, the major increase was at 60% time-compression.

Initial Position. Reference to Table 1 shows that the / ʃ/ and / tʃ / phonemes were perceived correctly most often (97% and 90%, respectively), whereas the voiced labio-dental / v / was perceived correctly least often (25%). The most common substitution for / v / was the voiced bilabial / b /. The remaining phonemes have percentage correct scores from 54% to 88%. The phonemes / p, t, g, v, h, and r, / revealed the highest degrees of consistent errors and suggest that most of the substitutions were ones associated with voicing, whereby a voiced phoneme was replaced by a voiceless phoneme and vice versa, i.e., b/f and g/hw. This was followed by errors associated with place of articulation (t/k and g/d). The b/v substitution, while not specifically

TABLE 2.--Phonemic confusions for phonemes in the final position over all conditions of time-compression at low sensation levels.

	p	b	t	d	k	g	f	v	θ	s	z	S	m	n	ŋ	tS	dz	r	l	NR	S %	C %
p	102	1	14	2		1	7		3	4				1		2			1	6	10	71
b		17	3					1	3				13	5						6	27	35
t	16	2	200	4	27	1	5	3	7	4		1		2		1			1	14		69
d	2		43	46	14	9		3	3	2	1	1	2	10					2	6	30	32
k	11	2	10		145		1	1	1	3				3						13		76
g			4	4	2	68	1			2		3	2	1	1		1	1		4		71
f	35	3	5		1	1	45			1						1	2	2		2	37	47
v			1	1		1		39						1				1	2	2		81
θ		1	1	3	4		7		26	1			1							4	15	54
s			8		4					124					1	2				4		86
z			7	1	2					5	70		1	5						3		73
S			2	1								40				2	3					83
m	7		1		3	2	2	4	1	1		1	85	7	1			1	3	9	13	58
n	2		13	21	9	2	2	6	4	2			12	105	4	1		1		14		63
ŋ			1	1	2								1		42					1		88
tS					4	2			1			8				75				6		78
dz				1	3					2				1		1	40					83
r			6		5									3		1		124		5		86
l			1	1	1	3	3							2				10	206	8		86

NR = No response; S = Substitution; C = Correct

fitting the analytic paradigm utilized, should be considered a place error, i.e., from a labio-dental to bilabial substitution.

Final Position. All phonemes appear to be less stable in the final position. Whereas the phoneme / ʃ / was perceived correctly 97% in the initial position, it was perceived correctly 83% in the final position. Further, no phonemes were perceived correctly 90% or more in the final position.

In reference to Table 2, the nasal / ŋ / was perceived correctly the largest percentage of time (88%). In contrast the voiced lingual-dental plosive / d / was perceived correctly least often (32%) and was replaced by its voiceless counterpart / t / 30% of the time. Further, the phoneme / g / was perceived correctly 35% of the time and was most often substituted with the bilabial nasal / m / (27%). The voiceless labial-dental fricative / f / was substituted by the voiceless plosive / p / 37% of the time and was perceived correctly 47%. Again, voicing and place of articulation appear to play a major role in determining substitution patterns.

It should be noted that in the initial position, the phoneme / v / had the lowest percentage correct score (25%), whereas in the final position it was correctly perceived 81% of the time.

The phonemes / t / and / p / tended to be substituted most often for other phonemes in the final position. This was not the case in the initial position where the phoneme / b / was substituted most often.

Combined Data of Initial and Final Phonemes. Table 3 shows the errors made when the initial and final positions are combined. The scores of this condition fell between those scores obtained when the initial and final positions were looked at separately. This becomes important when considering the total intelligibility of the / v / and / f / phoneme.

The sibilent phonemes appeared to be the most stable in the combined initial and final position. They were perceived correctly over 80% of the time under all conditions. In reference to Table 4, again it is the sibilents which are characterized by their durational aspects were perceived correctly the most often 85%. In contrast the voiceless labio-dental / f / was least stable (51%).

When viewing Table 4, the voiceless phonemes when considered together at low sensation levels were perceived correctly 73% of the time. The voiced and voiceless bilabial plosives / b / and / p / were most often substituted for the / f / phoneme (21% and 18% respectively). The phoneme / v / was correctly perceived

TABLE 3.--Phonemic confusions for phonemes in the initial and final positions combined over all conditions of time-compression at low sensation levels.

	p	b	t	d	k	g	f	v	θ	s	z	S	m	n	ŋ	tS	dz	r	j	h	l	hw	NR	S %	C %
p	205	6	18	3	16	2	13		3	4				1		2				1	1		13		71
b	2	121	3	5	3	10	2	1	3	1			14	5		1	1	1	1		2		16		63
t	19	2	296	9	45	2	7	3	7	5		1	1	2		1			3	4	1		23	k 10	69
d	2	6	43	159	16	13	1	3	3	2	1	1	2	12		1	3			1	3		17	t 15	55
k	15	4	20	10	238	6	2	2	2	3			2	4				4		1	1		22		71
g	2	1	4	25	4	133	1		1	2		3	3	1	1		2	3					6	d 13	69
f	35	41	5		1	2	97			2						1		2		1	4		4	b 21	51
v		20	2	2	1	1		51					6	2		1		1		1			4	b 21	53
θ		6	1	3	8	2	8		54	2			1						1	3	1		7		56
s		2	9	7	4					244		2				2	3	2		2			12		85
z		7	1	2	2					5	20		1	5									3		73
S			2	2								133				2	1	1			8		1		92
m	7	12	18	2	4	6	3		5	1		1	189	11	1			2		6	1		12		66
n	4	1	13	25	9	2	2	6	1	2			14	228	4		4	2					21		68
ŋ			1	1	2								1		42		124						1		88
tS		2	3		4		2		1			8				161							7		84
dz		1		5	4	2				2				1		2			1				2		86
r		1	6	2	5		1							3		1		23					12		84
j				1		3							1						41				2		85
h		1			1	19	1													117			5	g 13	81
l		8	1	3	3	3			4	2			3	2			1	13		3	368		18		85
hw				2		3							1					11		2		116	9		81

NR = No response; S = Substitution; C = Correct

TABLE 4.--Percentage of correct perception of distinctive features at the low sensation levels.

+8 and 16 dB SL	All Time-Compression Levels		
Feature	Number of Times Correctly Perceived	Total	- %
Nasal	459	672	68%
Voiced	1924	2640	73%
Unvoiced	1545	2112	73%
Duration	571	672	85%
Place of Articulation			
Front	663	1056	63%
Middle	1295	1824	68%
Back	1511	1872	81%

53% of the time and was most often substituted by the / b / phoneme. This is due to the small percentage obtained for the / v / phoneme in the initial position (25%). Although the voiceless phoneme / Θ / was perceived correctly only 53%, no clear substitution pattern was exhibited.

Further, Table 4 shows the phonemes classified into distinctive features and the percentage of correct perception associated with each classification. This table represents phonemes in both the initial and final position collapsed over all degrees of time-compression. Those phonemes classified under duration were perceived correctly most often (85%). At the low sensation levels, both the voiced and voiceless phonemes were perceived correctly 73% of the time. Table 4 reveals that those phonemes classified under nasality were perceived correctly 68% of the time.

Those phonemes whose point of articulation is the back of the oral cavity were perceived correctly most often (81%) when compared with those phonemes with middle and frontal points of articulation (middle, 68%, and front, 63%).

<u>High Sensation Level</u>

Tables 5 through 8 depict the results obtained when the sensation levels of 24 dB and 32 d B were considered across the several levels of time-compression

TABLE 5.--Phonemic confusions for phonemes in the initial position over all conditions of time-compression at high sensation levels.

	p	b	t	d	k	g	f	v	θ	s	z	S	m	n	tS	dz	r	j	h	l	hw	NR	S %	C %
p	130				5	4	1			1												3		90
b		135				5		2														2		94
t			131		7	1																5		91
d		1		138		1													4					96
k			13		131																		+ 9	91
g				10	1	83										1		1					d 10	87
f		5					90		1															94
v		4						31					10						1	2			M 20	65
θ		1				1			44													2		92
s										141												3		98
S												95										1		99
m		2											136				1		2			3		95
n			1	1									1	91								2		95
tS					1										95									99
dz				2	1											93								97
r		2											2				180		3	1		4		94
j																		48						100
h						1													133			2		92
l									1										6	176		9		92
hw																	1				142	1		99

NR = No response; S = Substitution; C = Correct

(0%, 30%, 40%, 50%, 60%, and 70%). It should be noted that most of the errors can be attributed directly to the time-compression procedure as the intensity of the signal is great enough for accurate intelligibility.

Tables 5 through 7 reveal those errors made when the phoneme was in the initial position, final position, and initial and final position combined. Also recorded in percentage scores is the percentage of times the phoneme was perceived correctly and the percentage of time it was perceived as another phoneme where applicable.

Comparison of Table 5 (initial errors) and Table 6 (final errors) revealed that overall there were more final phonemic confusions than initial position confusions. Further, reference to Tables 13 to 24 in Appendix A revealed that the major increase for initial position errors did not occur until the 70% time-compression level, whereas the majority of the final position errors did not occur until the 60% time-compression level.

The results further indicated that as the percentage of time-compression increases, intelligibility decreases (see Appendix A, Tables 13 to 24). In addition, it should be noted that the decrease in intelligibility is relatively gradual over the several conditions of time-compression until 70%, at which point there was a dramatic breakdown in overall intelligibility.

Initial Position. Table 5 reveals that all phonemes except / g / and / v / in the initial position were perceived correctly 90% (or better) of the time through all conditions of time-compression. The voiced gutteral / g / was correctly perceived 87% with the voiced lingual-dental / d / being substituted 10% of the time. The voiced labio-dental / v / was correctly perceived least often (65%) and the nasal bilabial / m / being substituted for it (20%). This suggests that the substitutions were ones associated with place of articulation. It should be noted that at the lower sensation levels (8 and 16 dB) the phoneme / b / was substituted for / v / whereas at the 24 dB and 32 dB sensation level, the phoneme / m / was substituted.

At the 70% time-compression level, phonemes in the initial position were dramatically reduced in intelligibility when compared with the other levels of time-compression (see Appendix A, Tables 13 through 24). The percentages of correct perception of the phonemes were all depressed when compared with Table 5. The phonemes / j / and / ʃ / tended to be the most stable at the 70% time-compression level (100% and 94% respectively). Further, definite substitution patterns occurred. The voiced gutteral phoneme / g / was substituted for the voiceless back phoneme / h / 38% of the time. It is interesting to note that the / k / phoneme was most often substituted for the voiceless lingual-dental plosive / t /, and vice versa (see Table 24 in Appendix A).

<u>Final Position</u>. Table 6 shows the results in the final position when compared with Table 5, where 18 out of 20 phonemes were perceived correctly 90% (or better) of the time in the initial position, Table 6 reveals that only 12 of 19 phonemes were correctly perceived 90% of the time or better. Dramatic decreases are seen in the / Θ / phoneme (92% in the initial position and 58% in the final position), the / d / phoneme (96% vs. 79%), the / b / phoneme (14% vs. 76%), and the / n / phoneme (95% vs. 79%). Table 6 reveals that the labio-dental / v / was correctly perceived 100% of the time in the final position, when compared with Table 5 the / v / phoneme was correctly perceived 65% of the time in the initial position over all levels of time-compression.

<u>Initial and Final Positions</u>. Table 7 shows the errors made when the initial and final positions are combined over all levels of time-compression. As can be seen all of the phonemes were perceived within the eighty and ninety percentage correct categories, except the / Θ / phoneme. The / Θ / phoneme was perceived correctly least often (75%) and was substituted most often by the voiceless labio-dental / f / phoneme (15%). Table 8 reveals

TABLE 6.--Phonemic confusions for phonemes in the final position over all conditions of time-compression at high sensation levels.

	p	b	t	d	k	g	f	v	θ	s	z	S	m	n	ŋ	tS	dz	r	l	NR	S %	C %
p	134	1	3			1	2													3		93
b		36											6	3					1	2	13 [m]	75
t	14		248		16			1	1	2				1						5		86
d		1	30	93	2	8				1				1					2	6	21 [t]	65
k			3		187		1											1				97
g			1		3	86							1	1						4		90
f	5		1				86							1					1	2		90
v								48														100
θ	1				2		14		28				1							2	29 [f]	58
s								1		143												99
z			2							4	90											94
S												47				1						98
m	4	1					2						133	1						3		92
n			3	25	1		1					1	15	189	1					1	10 [d]	79
ŋ															48							100
tS												2				91	1			2		95
dz																	47			1		98
r				2	2		1											139				97
l			2	1	1			1										2	231	2		96

NR = No response; S = Substitution; C = Correct

TABLE 7.--Phonemic confusions for phonemes in the initial and final positions combined over all conditions of time-compression at high sensation levels.

	p	b	t	d	k	g	f	v	θ	s	z	S	m	n	ŋ	tS	dz	r	j	h	l	hw	NR	S %	C %
p	264	1	3		5	5	3			1													6		92
b		171				5		2					6	3							1		4		89
t	14		379		23	1		1	1	2				1									10		88
d		2	30	231	2	9				1				5							2		6	t 13	80
k			16		318		1											1							95
g			1	10	4	169							1	1			1		1				4		88
f	5	5	1				176		1					1							1		2		92
v		4						79					10	1							2			m 10	82
θ	1	1			2	3	14		72				1										2	f 15	75
s								1		284													3		99
z			2							4	90														94
S												142				1							1		99
m	4	3					2						269	3				1					6		93
n			4	26	1	1	1		1	1			16	280		1							4		83
ŋ															48										100
tS					1							2				186	1						2		97
dz				2	1												140						1		97
r		2		2	2			1					2					319		3	1		4		95
j																			48						100
h						9														133					93
l			2	1	1			1	1									2		6	407				92
hw																		1				142	1		99

that this substitution is apparently related to a place of articulation type error. It should also be noted that Table 8 reveals that all phonemes classified under frontal place of articulation including the / f / phoneme, a score of 91% was obtained for intelligibility. In contrast for those phonemes classified as having middle oral cavity place of articulation, including the / θ / phoneme, a score of 89% was obtained for intelligibility.

Further, Table 8 shows that phonemes classified under duration were perceived correctly most often (98%). In contrast those phonemes classified under nasality and those under middle place of articulation were perceived correctly least often (89%). Those scores were obtained for phonemes in both the initial and final positions under all levels of time-compression.

TABLE 8.--Percentage of correct perception of distinctive features at the high sensation levels.

+24 and 32 dB SL	All Time-Compression Levels		
Feature	Number of Times Correctly Perceived	Total	- %
Nasal	597	672	89%
Voiced	2393	2640	91%
Unvoiced	2140	2304	93%
Duration	656	672	98%
Place of Articulation			
Front	1618	1824	89%
Middle	959	1056	91%
Back	1770	1872	95%

CHAPTER IV

DISCUSSION

Many investigators have considered distorted speech tests as a means of distinguishing lesions of the higher auditory pathways from those involving the cochlea (Bocca and Calearo, 1963; Bocca, 1967). One form of distorted speech is electro-mechanical time-compression of the speech signal as developed by Fairbanks, Everitt, and Jaeger in 1954. The potential use of time-compressed speech for detecting lesions of the higher auditory pathway has been demonstrated by Calearo and Lazzaroni (1957); deQuiros (1964), and Beasley, Schwimmer, and Rintelmann (1972). In the latter study, normative data were gathered using a standardized discrimination test (the Northwestern Auditory Test No. 6). To date no analysis has been undertaken to determine what effects the time-compression procedure has on phonemic perceptual confusions. A discussion of these effects is relevant to further research involving the time-compression procedure.

Consistency of Phonemic Errors

This study has shown that the replacement of one phoneme by another is not a random process, a finding supported by Lafon (1961) and Miller and Nicely (1955). This study further revealed that the intensity level of presentation above speech threshold played a role in the intelligibility of the stimuli. This occurred for all conditions of time-compression. The effects of intensity increments lessen, as would be expected, as an optimal listening intensity is approached. Calearo and Lazzaroni (1957) noted that there is a tendency for intensity to neutralize the effects of speech acceleration for normal subjects. This was further supported by Beasley et al. up to the 70% time-compression ratio, whereby a dramatic decrease in intelligibility occurred regardless of sensation level.

The consistency of the errors made is also altered by the level of presentation. At low sensation levels the phonemes most affected in the initial position were / p, t, k, f, Θ, g, and v /. At the high sensation levels, only the initial position phonemes / g / and / v / fell below a 90% correct intelligibility score through the first five levels of time-compression. At the 70% compression level not only did the / g / and / v / remain unintelligible, but in addition the initial position phonemes / p, t, k, f, Θ, l, and h / became unintelligible. This was also true at 70% for low sensation levels of presentation. It should be noted that many of the errors made at the low sensation levels may be due to the low intensity of presentation as well as to the varying time-compression ratios.

Those errors made in the final position do not necessarily correspond to those made in the initial position. An example of this is the / v / phoneme whose percentage of intelligibility is low in the initial position, but becomes much more intelligible in the final position. These results were displayed at both low and high sensation levels of presentation. Also, in the initial position the / v / phoneme was perceived as / b / at low sensation levels. This is in support of Lafon (1961) who also found that the / b / phoneme was most often substituted for / v /. In contrast at the higher sensation levels, the phoneme / v / was most often substituted by the / m / phoneme. It must be noted that the / m / for / v / substitution occurred at the 70% time-compression ratio (see Appendix A, Table 24). This phenomenon might be explained through the durational aspects of the / m / phoneme. The / m / phoneme is longer in duration than either the / v / or / b / phonemes (Miller and Nicely, 1955). Consequently, at the 70% time-compression level, neither the / v / or / b / phoneme are long enough in duration to be temporally processed, so the / m / phoneme, which is longer in duration and has similar voicing and place characteristics as / v /, is confused for the / v / phoneme. Further, the / m / phoneme is also substituted for / v / at the lower sensation levels, but not to the extent that the / b / phoneme is substituted (see Appendix A, Tables 1 through 24).

Initial and Final Positions

This study revealed that more errors occurred in the final position than in the initial position. This held true regardless of the level of presentation. Whereas the 70% time-compression level was the point at which phonemes in the initial position became most affected, in the final position the intelligibility of the phoneme was affected at both the 60% and 70% time-compression level. Lehiste (1964) and Hoard (1966) suggested that consonants near the beginning of a syllable tend to be longer than they are near the end. In contrast, Barnwell III (1970) found this not to be the case for stop consonants. He found no measurable differences for the stop consonants in the initial and final positions. It may be speculated that the durational aspects of the final phoneme is shorter than that of the initial phoneme in conversational speech. Then when the signal is temporally processed the point of reaching the minimal time needed for temporal processing occurs at a lower level of compression than it does for phonemes in the initial position.

Effects of Time-Compression Levels

Although the effects of time-compression ratios and levels of presentation cannot be separated when discussing their effects upon intelligibility of phonemes, it is felt that the role that time-compression plays in determining intelligibility can be discussed for the +24 and 32 dB sensation levels. The results of this study indicate that a gradual decrease occurred in phonemic intelligibility as the duration of the signal was reduced. This decrease was gradual until the 70% time-compression level, where a dramatic reduction in intelligibility occurred. It is felt that these errors are due directly to the time-compression procedure and deserves some discussion. Table 9 shows those phonemes most affected and those phonemes most often substituted for them.

Many of these substitutions can be explained according to their durational characteristics. For example, the voiceless phoneme / p / consumes less time in production (Jakobson et al., 1952) than voiced plosives such as the / g / which has been substituted for / p /. The longer durational aspects of the voiced plosive / g / allows more time for temporal processing.

TABLE 9.--Phonemes most frequently missed at the 70% time-compression level in the initial position. Also shown are those phonemes most often substituted for them and the percentage of substitution.

Phoneme	% Correct	Phoneme Substituted	%
p	54%	g	17%
t	63%	k	20%
k	87%	t	13%
f	69%	b	25%
b	75%	g	17%
g	50%	d	31%
v	13%	m	50%
l	50%	h	19%
h	54%	g	38%

Thus, the intelligibility of the sound is increased and enhances its substitution for other phonemes with shorter durational characteristics, such as / p /. Further, the / g / phoneme's place of articulation plays a role in its substitution for / p /. This study has shown, those phonemes whose place of articulation are the back of the oral cavity (such as / g /) are more intelligible than those phonemes whose place of articulation are either the middle or front of the oral cavity. Thus, the / g / phoneme is more intelligible because of its place of articulation and longer durational characteristics and is more likely to be substituted for other phonemes. This is further demonstrated by the / g / for / b / substitution and / g / for / h / substitution patterns. Lafon (1961) reported that the phoneme / k / is most often substituted for / t /, and was further supported by this study. This substitution may be explained by discussing the characteristic qualities of the two phonemes. Both phonemes are voiceless plosives and only differ in their place of articulation and timbre, thus making the distinction between the sounds much more difficult for the listener and enhancing the substitution pattern exhibited.

In contrast, in the final position, the / t / phoneme is replaced by the / p / phoneme and the / b / phoneme is replaced by the / m / and / n / phonemes.

These changes in substitutions may be explained by the role which the vowel plays upon the phoneme following it. This is supported by Lehiste and Shockey (1971) who found that the vowel carries coarticulatory information as to the identity of the following consonant.

Distinctive Features

The role of distinctive features upon the intelligibility of the phoneme is important in aiding the listener to distinguish one phoneme from another. In order to decode the message, the receiver extracts the distinctive features from the perceptual data (Jakobson, Fant, and Halle, 1963). Miller and Nicely (1955) confirmed that the perception of each of these features is relatively independent from the perception of the others.

In this study the phonemes were classified into five groups of distinctive features:

(1) voicing, which included the phonemes / b, d, g, v, z, ʤ , m, n, ŋ , l, r, hw, and j /

(2) unvoicing, / p, t, k, f, θ, s, ʃ, tʃ, and h /

(3) nasality, / m, n, and ŋ , /

(4) duration, / s, ʃ, ʤ , and z, / and

(5) place of articulation, which was divided into three subgroups,

(a) front, which included the phonemes / p, b, f, v, and m /

(b) middle, / t, d, θ, s, z, n, ŋ, j, and tʃ / and

(c) back, / k, g, ʃ, dʒ, r, l, h, and hw /.

Miller and Nicely (1955) examined 16 initial consonants followed by the vowel / a / using varying signal-to-voice ratios and band-pass filters. The results revealed that voicing and nasality were much less affected by a random mashing noise. In contrast, this study examined consonants in both the initial and final positions with various vowels. The results of this study revealed that the phonemes characterized by the distinctive features of nasality and voicing were also affected by time-compression along with those phonemes characterized by other features. This occurred at both high and low sensation levels. Further, it showed that those phonemes with the greatest duration were least affected by time-compression. This finding is expected and supported by Jakobson (1952) in his discussion that phonemes with the greatest duration are most intelligible.

Conclusions

In summary, then, this study has shown that a consistency in phonemic errors does exist and may be due to a change in the speech signal by various degrees of time-compression. Further, these phonemic confusions differ depending upon the placement of the phoneme in the word. Different substitution patterns were exhibited for the same phoneme when in the initial position of the word as opposed to the final position.

The sensation level of presentation also affected the type of phonemic errors. An example of this was the / b / for / v / substitution at low sensation levels whereas the / m / phoneme was substituted for the / v / phoneme at the high sensation levels.

Finally, this study has shown that the distinctive features of the phonemes were also affected by the time-compression procedure. Those phonemes with the greatest duration tended to be least affected by the various time-compression ratios.

Implications for Further Research

A spectographic study of the phonemes would help to determine what effect the time-compression procedure has on their physical characteristics. A study of this nature might also be beneficial in further helping to explain many of the substitution patterns that were demonstrated in this study. With a spectrographic study, it might be determined if the energy and/or formant structure of the phonemes are altered by time-compression, causing one phoneme to take on physical characteristics of another phoneme, thus enhancing the substitution patterns exhibited in this study.

Perhaps the most obvious area for further research is to analyze lists II, III, and IV of Form B of the NU-6 in the same manner. The comparison can then be made between those results and the results obtained in this study. Beasley et al. (1972) found list I to be the most difficult of these lists, and list IV to be the least difficult. Thus, a comparison of list I and list IV may be beneficial in confirming their degree of difficulty. Another advantage of comparing all four lists would be in averaging those scores obtained in order to further substantiate the findings of this study.

Another area of research is the effect time-compression has on vowels. This should also be done for all four lists of Form B of the NU-6. With this type of study it could be determined if vowels, like consonants, display a definite substitution pattern. Further, it would determine which vowels retain their intelligibility during the time-compression procedure.

LIST OF REFERENCES

LIST OF REFERENCES

Aaronson, D., Temporal factors in perception and short term memory. Psychological Bulletin 130-144 (1967).

Barnwell III, T. P., Some syllabic junctural effects in English. MII Quarterly Progress Report 99, 149-159 (1970).

Beasley, D. S., Schwimmer, S., Rintelmann, W. F., Intelligibility of Time-Compressed CNC Monosyllables, Journal of Speech and Hearing Research, in press (1972).

Beasley, D. S., and Shriner, T. H., Auditory Analysis of temporally distorted sentential approximations. Audiology: Journal of Auditory Communication, in press (1972).

Black, J. W., Journal of Speech and Hearing Disorders, 17, 409 (1952).

Black, J. W., and Singh, S., The psychological bases of phonetics, Manual of Phonetics, (ed.) B. Malmberg Netherlands, 105-125 (1970).

Bocca, E. and Calearo, C., Central hearing processes. Modern Developments in Audiology. J. Jerger, ed., New York: Academic Press (1963).

Bordley, J. E., and Haskins, H. L., The role of the cerebrum in hearing. Annals of Otology, Rhinology, and Laryngology, LXIV, 370-382 (1955).

Calearo, C., and Lazzaroni, A., Speech intelligibility in relation to the speed of the message. Laryngoscope, 67, 410-419 (1957).

Daniloff, R. G., Shriner, T. H., and Zemlin, W. R., Intelligibility of vowels altered in duration and frequency. Journal of the Acoustical Society of America, 44, 700-707 (1968).

deQuiros, J., Accelerated speech audiometry, and examination of test results (Trans. by J. Tonndory). Translations Beltone Institute of Hearing Research, No. 17, Beltone Institute of Hearing Research: Chicago (1964).

Eldert, E., and Davis, H., The articulation function of patients with conductive deafness. Laryngoscope, 61, 891-909 (1951).

Fairbank, G., Everitt, W., and Jaeger, R., Methods for time of frequency compression--expansion of speech. Translations I. R. E. P. G. A., AU-7, 7 - 12 (1954).

Fletcher, H., Speech and Hearing in Communication. Van Nostrand Co., New York (1958).

Fournier, J. E., L' analyse et l' identification du message sonore. Journal Franc. Oto-Rhino-Laryngol., 67 (1956).

Hirsh, I. J., Davis, H., Silverman, S. R., Reynolds, E. G., Eldert, E., and Benson, A. W., Development of material for speech audiometry. Journal of Speech and Hearing Disorders, XVI, 321-328 (1952).

Hoard, J. E., Juncture and syllable structure in English, Phonetica, 15, 96-109 (1966).

House, A. S., and Fairbanks, g., The influence of consonant environment upon the secondary acoustical characteristics of vowels. Journal of the Acoustical Society of America, 25, 105-113, (1953).

Jakobson, A., Fant, C. G. M., and Halle, M., Preliminaries to speech analysis. Mass. Inst. of Technology, Acoustic Laboratory, Technical Report, No. 13, (1952).

Jakobson, R., and Halle, M., Phonology in relation to phonetics. Manual of Phonetics, Malmberg, ed. Netherlands (1970).

Katz, J., Differential diagnosis of auditory impairments. Audiometry for the Retarded. Robert T. Fulton and Lyle L. Lloyd, eds. Baltimore: Williams and Wilkins Company (1969).

Lafon, J. C., Message et phonetique. Presses Univ. de France, Paris, (1961).

Lehiste, I., Juncture. Proc. of the Fifth International Congress of Phonetic Sciences (1964).

Lehiste, I., and Peterson, G., Linguistic considerations in the study of speech intelligibility. Journal of the Acoustical Society of America, 31, 280-286 (1959).

Lehiste, I., and Shockey, L., Coarticulation effects in the identification of final plosives, paper presented at the 82nd meeting of the Acoustic Society of America, Denver, Colorado, October 19-22 (1971).

Malmberg, B., Manual of Phonetics. Bertil Malmberg, ed. Netherlands (1970).

Miller, G. A., Heise, G. A., and Lichten, W., The intelligibility of speech as a function of the context of test material. Journal of Experimental Psychology, 41, 229-335 (1951).

Miller, G. A., and Nicely, P. E., An analysis of perceptual confusions among some English consonants. Journal of the Acoustical Society of America, 27, 338-352 (1954).

Peterson, G. E., and Borney, H. L., Control methods used in a study of the vowels. The Journal of the Acoustical Society of America, 24, 175-184 (1952).

Rintelmann, W. F., and Jetty, A. J., Reliability of speech discrimination testing using CNC monosyllabic words. Unpublished study, Michigan State University (1968).

Shriner, T. H., Beasley, D. S., and Zemlin, W. R., The effects of frequency division on speech identification in children. Journal of Speech and Hearing Research, 12, No. 2, 413-422 (1969).

Stevens, S. S., Handbook of Experimental Psychology. John Wiley and Sons, Inc. New York: (1948).

Thorndike, E. L., and Lorge, I., Teacher's Word Book of 30,000 Words. New York: Bureau of Publications, Teacher's College, Columbia University (1950).

Tillman, T. W., and Carhart, R., An expanded test for speech discrimination using CNC monosyllabic words: Northwestern University Auditory Test No. 6 (1966).

Tillman, T. W., Carhart, P., and Wilbur, L. A., United States Air Force School of Aerospace Medicine, SAM-TDR-62-135, AD No. 403275, Brooks Air Force Base, January, 1963.

Williford, J. A., Audiological evaluation of central auditory disorders. Maico Audiological Library Series, VI, 1-7 (1969).

Zemlin, W. R., Daniloff, R. G., and Shriner, T. H., The Difficulty of listening to time-compressed speech. Journal of Speech and Hearing Research, 11, 875-881 (1968).

APPENDICES

APPENDIX A

CONFUSION MATRICES FOR PHONEMES IN THE INITIAL AND FINAL POSITIONS AT EACH LEVEL OF TIME-COMPRESSION FOR LOW (8 and 16 dB) AND HIGH (24 and 32 dB) SENSATION LEVELS

TABLE A-1.--Phonemic confusions for phonemes in the initial position at 0% time-compression at low sensation levels.

	p	b	t	d	k	g	f	v	θ	s	z	S	m	n	tS	dz	r	j	h	l	hw	NR	S %	C %
p	20		2		1		1																	83
b		20				2							1							1				83
t			22	1															1					92
d				24																				100
k	1		1		21														1					88
g				1		13										1		1						81
f		2					13													1				81
v		2						4					1	1									25	50
θ									7	1														88
s										24														100
S												16												100
m							1						22							1				92
n														16										100
tS															16									100
dz																16								100
r																	32							100
j																		8						100
h						1													23					96
l																				32				100
hw																					24			100

NR = No response; S = Substitution; C = Correct

TABLE A-2.--Phonemic confusions for phonemes in the final position at 0% time-compression at low sensation levels.

	p	b	t	d	k	g	f	v	θ	s	z	S	m	n	ŋ	tS	dz	r	l	NR	S %	C %
p	19		1				1		1	1				1								79
b		6											2								25	75
t		1	37	2	5		2							1								77
d				20										4								83
k	2	1			29																	90
g						14						2										88
f	1	2					13															81
v								8														100
θ				1	1		1		5													63
s										24												100
z				1							15											94
S												7				1						88
m			1										21	1					1			88
n				2									2	36								90
ŋ					1										7							88
tS												1				15						94
dz																	8					100
r																		24				100
l						1								1					38			95

NR = No response; S = Substitution; C = Correct

TABLE A-3.--Phonemic confusions for phonemes in the initial position at 30% time-compression at low sensation levels.

	p	b	t	d	k	g	f	v	θ	s	z	S	m	n	tS	dz	r	j	h	l	hw	NR	S %	C %
p	22				1														1					92
b		21					1									1		1						88
t	1		16	3	2					1									1					67
d		2		21															1					88
k		1	3	2	17								1											71
g		1				14												1						88
f		4					12																25	75
v		6						2															75	25
θ		1							6									1						75
s										24														100
S												16												100
m		1		1									19						2	1				79
n		1												15										94
tS															16									100
dz																15		1						94
r																	29			3				91
j																		8						100
h						2													22					92
l																				32				100
hw																	1		2		21			88

NR = No response; S = Substitution; C = Correct

TABLE A-4.--Phonemic confusions for phonemes in the final position at 30% time-compression at low sensation levels.

	p	b	t	d	k	g	f	v	Θ	s	z	S	m	n	ŋ	tS	dz	r	l	NR	S %	C %
p	17		4	1						1						1					17	71
b		11							2				2								25	50
t	2		40		3			1	1										1			84
d			1	14	1	5							1						2		21	58
k	2		1	27						1				1								84
g					1	14											1					88
f	7						9														44	56
v								8														100
Θ									7	1												88
s			3							21												88
z				1							14			1								88
S			1									7										88
m				1				1				1	15	4				1	1		17	63
n			2	1									4	33								83
ŋ															8							100
tS							2					2				12						75
dz				1										1			6					75
r														1				23				96
l	1								1									1	37			93

NR = No response; S = Substitution; C = Correct

TABLE A-5.--Phonemic confusions for phonemes in the initial position at 40% time-compression at low sensation levels.

	p	b	t	d	k	g	f	v	θ	s	z	S	m	n	tS	dz	r	j	h	l	hw	NR	S %	C %
p	19		1		1		3																f 12	79
b		20		2		2																		83
t			18		2	2	1												1					75
d		3		20																1				83
k			2		21								1											87
g				2		14																		87
f		6					10																b 37	62
v		2		1	1			1					2							1			hcm 12	12
θ		2							6														b 25	75
s		1	1	1						20		1												83
S												16												100
m		3				1							18	1					1				b 12	75
n													1	14						1				87
tS															14	2								87
dz																16								100
r																	25			7			l 22	78
j				1		2												5						62
h						6													18				g 25	75
l										1										31				96
hw				1													1				22			91

NR = No response; S = Substitution; C = Correct

TABLE A-6.--Phonemic confusions for phonemes in the final position at 40% time-compression at low sensation levels.

	p	b	t	d	k	g	f	v	Θ	s	z	S	m	n	ŋ	tS	dz	r	l	NR	S %	C %
p	20		3						1												12 (t)	83
b		3	2					1					1	1							25 (t)	37
t	2		37		6				2					1								77
d			10	7	2	1		1	1					2							41 (t)	29
k	2		3		25		1		1													78
g				2		10	1			2					1							62
f	4		1				10			1											25 (p)	62
v								8														100
Θ		1			3	1			3												37 (k)	37
s		1	1							21						1						87
z			2								13		1									81
S												8										100
m			2				2						18	1					1			75
n	1		2	4	1		1						4	25	1	1						62
ŋ					2										6						25 (k)	75
tS					1				1							14						87
dz																	8					100
r														2				22				91
l						1												1	38			95

NR = No response; S = Substitution; C = Correct

TABLE A-7.--Phonemic confusions for phonemes in the initial position at 50% time-compression at low sensation levels.

	p	b	t	d	k	g	f	v	θ	s	z	S	m	n	tS	dz	r	j	h	l	hw	NR	S %	C %
p	19	1			3		1																	79
b	1	18		1	3									1										75
t	1		17		4								1						1				16 (k)	70
d				22										2										91
k	1				19	2											2							79
g				8		7							1										50 (d)	43
f		9					7																56 (b)	43
v		3						3					1						1				37 (b)	37
θ									5										3					62
s		1		1						22														91
S												16												100
m													21	1					1	1				87
n				2										14										87
tS															16									100
dz																16								100
r																	31			1				96
j																		8						100
h							1												23					95
l		1															2			29				90
hw							1										7				16		29 (r)	66

NR = No response; S = Substitution; C = Correct

TABLE A-8.--Phonemic confusions for phonemes in the final position at 50% time-compression at low sensation levels.

	p	b	t	d	k	g	f	v	θ	s	z	S	m	n	ŋ	tS	dz	r	l	NR	S %	C %
p	19		1				2		1	1												79
b		2							1				4	1							m 50	25
t	3		35		4		1		3	2												72
d	1		11	5	1	1			1	1	1			2							d 20	45
k	1		5		25					1											t 15	78
g						15							1									93
f	8					1	6									1					p 50	37
v								8														100
θ						1	1		6													75
s										24												100
z										1	12			2					1			75
S												7				1						87
m			3		1		1	1		1			16	1								66
n			2	5	1	1		4		1				25				1				62
ŋ															8							100
tS					1											15						93
dz																	8					100
r																		24				100
l													1					1	38			95

NR = No response; S = Substitution; C = Correct

TABLE A-9.--Phonemic confusions for phonemes in the initial position at 60% time-compression at low sensation levels.

	p	b	t	d	k	g	f	v	θ	s	z	S	m	n	tS	dz	r	j	h	l	hw	NR	S %	C %
p	19		1		3																	1		79
b		16		2		3														1		2		66
t	1		14		3	2												2				2		58
d		1		19													1					3		79
k		1	3	4	10		1		1													4	d 16	41
g						13										1						2		81
f		9					6			1													b 56	37
v		4						1					2									1	b 50	12
θ		1				4	1		2														g 50	25
s				1						19												4		79
S												16												100
m		2							1				15	2			1					3		62
n				1										14								1		87
tS															15	1								93
dz				2		1										13								81
r				1													25			4		2		78
j																		8						100
h						3													19			2		79
l																	1		2	24		5		75
hw						1											1				20	2		83

NR = No response; S = Substitution; C = Correct

TABLE A-10.--Phonemic confusions for phonemes in the final position at 60% time-compression at low sensation levels.

	p	b	t	d	k	g	f	v	θ	s	z	S	m	n	ŋ	tS	dz	r	l	NR	S %	C %
p	16	1		3			1			2										1		66
b		2											4							2	50	25
t	6	1	29	1	5			1								1				4	12	60
d			12	0	5			2	2					1						2	50	0
k	1		1		25					1										4		78
g			1	1	1	11							1							1		68
f	7		3				6														43	37
v				1				6												1		75
θ							3		4											1	37	50
s			1		1					21										1		87
z			3							2	9									2		56
S												8										100
m	4		6		2			2					8							2	25	33
n	1		4	6	2			2					2	17	1					5	15	42
ŋ			1											6						1		75
tS												3				12				1		75
dz																	8					100
r			4		2													18				75
l									2					1				2	33	2		82

NR = No response; S = Substitution; C = Correct

TABLE A-11.--Phonemic confusions for phonemes in the initial position at 70% time-compression at low sensation levels.

	p	b	t	d	k	g	f	v	θ	s	z	S	m	n	tS	dz	r	j	h	l	hw	NR	S %	C %
p	5	4		1	7	4	1															2	K 29	20
b	1	9				3	1			1							1					8	NR 33	37
t			9	2	6		1									1						5	K 25	37
d				7	1	4	1								1		2					8	NR 33	29
k	1		1	4	5	4								1			2		1			5	NR 20	20
g				8	2	4			1													1	d 50	25
f		9				1	4															2	b 56	25
v		3	1					1									1			1		1	b 37	12
θ		1		1	1	1			2													2		25
s				1	2					11			1				2		1	1		5		45
S				1								13					1					1		81
m		6		1	1	3							9						1	2		1	b 25	37
n		2	1										1	5			1					6	NR 37	31
tS		1	2												10	1				1		1		62
dz		1		2	1	1									1	8						2		50
r		1		1			1					1					16		7			5	L 21	50
j						1							1					4				2		50
h		1			1	6	1												12			3	g 25	50
l		7		2					1					1		1			1	14		5	b 21	43
hw				1		1							1				1				13	7	NR 29	54

NR = No response; S = Substitution; C = Correct

TABLE A-12.--Phonemic confusions for phonemes in the final position at 70% time-compression at low sensation levels.

	p	b	t	d	k	g	f	v	θ	s	z	S	m	n	ŋ	tS	dz	r	l	NR	S %	C %
p	11		2	1		1	3									1			1	4		45
b		0	1											3						4	NR 50	0
t	4	1	20	1	4	1	2	1	1	2		1								10	NR 20	41
d	1		7	0	5	2				1			1	1		1		1		4	T 29	0
k	3	1	2		14			1						2						9	NR 28	43
g	2		2	1		4								1	1					5		25
f	7	1	1		1		1											2		3	P 43	6
v			1			1		1						1				1	2	1		12
θ			1		2		2		0				1							2		0
s			3		3	1				13					1					3		54
z			1		2	2				2	7			1						1		43
S			1	1								3					3				dʒ 37	37
m	3		3		1	2							7		1				1	6		29
n			5	1	5	1	1		1	1				14	2					9	NR 22	35
ŋ													1		5					2		62
tS					2							2			7					5	NR 31	43
dz					3					2					1	2					K 37	25
r			2		3										1			13		5		54
l			1	1	1	1	3		1									4	22	6		55

NR = No response; S = Substitution; C = Correct

TABLE A-13.--Phonemic confusions for phonemes in the initial position at 0% time-compression at high sensation levels.

	p	b	t	d	k	g	f	v	θ	s	z	S	m	n	tS	dz	r	j	h	l	hw	NR	S %	C %
p	24																							100
b		24																						100
t			24																					100
d				24																				100
k					24																			100
g						16																		100
f							16																	100
v								8																100
θ									8															100
s										24														100
S												16												100
m													24											100
n														16										100
tS															16									100
dz																16								100
r																	32							100
j																		8						100
h																			24					100
l																				32				100
hw																					24			100

NR = No response; S = Substitution; C = Correct

TABLE A-14.--Phonemic confusions for phonemes in the final position at 0% time-compression at high sensation levels.

	p	b	t	d	k	g	f	v	θ	s	z	S	m	n	ŋ	tS	dz	r	l	NR	S %	C %
p	24																					100
b		8																				100
t			47		1																	97
d			1	21	1																	87
k					32																	100
g						16																100
f	1						15															93
v								8														100
θ						2	1		5													62
s								1		23												95
z											16											100
S												8										100
m													24									100
n														40								100
ŋ															8							100
tS												1				14	1					87
dz																	8					100
r																		24				100
l																			40			100

NR = No response; S = Substitution; C = Correct

TABLE A-15.--Phonemic confusions for phonemes in the initial position at 30% time-compression at high sensation levels.

	p	b	t	d	k	g	f	v	θ	s	z	S	m	n	tS	dz	r	j	h	l	hw	NR	S %	C %
p	23																							95
b		24																						100
t			24																					100
d				24																				100
k			1		23																			95
g						16																		100
f		1					15																	93
v		2						5															25	62
θ									8															100
s										24														100
S												16												100
m													24											100
n														16										100
tS															16									100
dz																16								100
r																	32							100
j																		8						100
h																			24					100
l																				32				100
hw																					24			100

NR = No response; S = Substitution; C = Correct

TABLE A-16.--Phonemic confusions for phonemes in the final position at 30% time-compression at high sensation levels.

	p	b	t	d	k	g	f	v	θ	s	z	S	m	n	ŋ	tS	dz	r	l	NR	S %	C %
p	24																					100
b		7											1									87
t			47		1																	97
d				21		3																87
k					31		1															96
g						16																100
f							15												1			93
v								8														100
θ									8													100
s										24												100
z											16											100
S												8										100
m													24									100
n					3								2	35								87
ŋ															8							100
tS																16						100
dz																	8					100
r																		24				100
l																			40			100

NR = No response; S = Substitution; C = Correct

TABLE A-17.--Phonemic confusions for phonemes in the initial position at 40% time-compression at high sensation levels.

	p	b	t	d	k	g	f	v	θ	s	z	S	m	n	tS	dz	r	j	h	l	hw	NR	S %	C %
p	24																							100
b		23																						96
t			22	2																				92
d				24																				100
k					24																			100
g		1				14											1							88
f							16																	100
v		1						3					4										50	38
θ									8															100
s										24														100
S												16												100
m													24											100
n														16										100
tS															16									100
dz																16								100
r																	32							100
j																		8						100
h																			24					100
l																				32				100
hw																					24			100

NR = No response; S = Substitution; C = Correct

TABLE A-18.--Phonemic confusions for phonemes in the final position at 40% time-compression at high sensation levels.

	p	b	t	d	k	g	f	v	θ	s	z	S	m	n	ŋ	tS	dz	r	l	NR	S %	C %
p	24																					100
b		8																				100
t	1		43	4																		89
d			4	18	1																16†	75
k					32																	100
g						16																100
f	1						15															93
v								8														100
θ	1						2		5													62
s										24												100
z											16											100
S												8										100
m													24									100
n				4									1	34	1							85
ŋ															8							100
tS												1				15						93
dz																	8					100
r																		24				100
l																			40			100

NR = No response; S = Substitution; C = Correct

TABLE A-19.--Phonemic confusions for phonemes in the initial position at 50% time-compression at high sensation levels.

	p	b	t	d	k	g	f	v	θ	s	z	S	m	n	tS	dz	r	j	h	l	hw	NR	S %	C %
p	24																							100
b		23						1																95
t			24																					100
d				23										1										95
k			4		20																		†16	83
g				2		14																		87
f							16																	100
v								8																100
θ									8															100
s										24														100
S												16												100
m													23	1										95
n														16										100
tS															16									100
dz																16								100
r																	32							100
j																		8						100
h																			24					100
l																				32				100
hw																					24			100

NR = No response; S = Substitution; C = Correct

TABLE A-20.--Phonemic confusions for phonemes in the final position at 50% time-compression at high sensation levels.

	p	b	t	d	k	g	f	v	Θ	s	z	S	m	n	ŋ	tS	dz	r	l	NR	S %	C %
p	22	1					1															91
b		6											2								25	75
t	1		43		4																	89
d			4	20																		83
k					32																	100
g						15							1									93
f	2						14															87
v								8														100
Θ							2		6												25	75
s										24												100
z			2								14											87
S												8										100
m													23	1								95
n				4										36								90
ŋ															8							100
tS																16						100
dz																	8					100
r																		24				100
l																			40			100

NR = No response; S = Substitution; C = Correct

TABLE A-21.--Phonemic confusions for phonemes in the initial position at 60% time-compression at high sensation levels.

	p	b	t	d	k	g	f	v	θ	s	z	S	m	n	tS	dz	r	j	h	l	hw	NR	S %	C %
p	22		2																					91
b		23				1																		95
t			22		2																			91
d				24																				100
k			5		19																		20	79
g				2		14																		87
f							16																	100
v								6				1								1				75
θ		1							7															87
s										24														100
S												16												100
m													23	1										95
n														16										100
tS															16									100
dz																16								100
r																	32							100
j																		8						100
h																			24					100
l																				32				100
hw																					24			100

NR = No response; S = Substitution; C = Correct

TABLE A-22.--Phonemic confusions for phonemes in the final position at 60% time-compression at high sensation levels.

	p	b	t	d	k	g	f	v	θ	s	z	S	m	n	ŋ	tS	dz	r	l	NR	S %	C %
p	21		1			1	1															87
b		6											1	1								75
t	7		36		4			1													14	75
d			8	12		4															33	50
k					32																	100
g						16																100
f			1				15															93
v								8														100
θ							6		1													75
s										24												100
z											16											100
S												8										100
m													24									100
n				4		1							3	36								81
ŋ															8							100
tS																16						100
dz																	8					100
r																		24				100
l			1		1													1	37			92

NR = No response; S = Substitution; C = Correct

TABLE A-23.--Phonemic confusions for phonemes in the initial position at 70% time-compression at high sensation levels.

	p	b	t	d	k	g	f	v	θ	s	z	S	m	n	tS	dz	r	j	h	l	hw	NR	S %	C %
p	13				3	4				1												3	g 17	54
b		18				4																2	g 17	75
t			15		5	1																3	k 20	63
d		1		19		1								3									n 13	79
k			3		21																		t 13	87
g				5	1	8																1	d 31	50
f		4					11		1														b 25	69
v		1						1					4	1						1			m 50	13
θ						1			5													2		63
s										21												3		88
S												15										1		94
m		2											18									3		75
n			2										1	11								2		69
tS				1											15									94
dz			2	1												13								81
r		2											2				20		3	1		4		63
j																		8						100
h						9													13			2	g 38	54
l									1										6	16		9	h 19	50
hw																	1				22	1		92

NR = No response; S = Substitution; C = Correct

TABLE A-24.--Phonemic confusions for phonemes in the final position at 70% time-compression at high sensation levels.

	p	b	t	d	k	g	f	v	θ	s	z	S	m	n	ŋ	tS	dz	r	l	NR	S %	C %
p	19		2																	3		79
b		1											2	2				1		2		
t	5		32		1				1	2				1						6		66
d			1	13	1	1													2	6		54
k			3		27													1		1		84
g			1		3	7								1						4	18 [K]	43
f	1						12							1						2		75
v								8														100
θ					2		3		3												37 [f]	37
s										24												100
z										4	12										25 [ʃ]	75
S												7				1						87
m	4		1					2					14							3	16 [p]	58
n			3	6	1		1		1	1			5	20						2	15 [d]	50
ŋ															8							100
tS																14				2		87
dz																	7			1		87
r				2	2				1									19				79
l								1											37	2		92

NR = No response; S = Substitution; C = Correct

APPENDIX B

LIST I, FORM B, NORTHWESTERN UNIVERSITY AUDITORY TEST, NUMBER 6

APPENDIX B

LIST I, FORM B, NORTHWESTERN UNIVERSITY

AUDITORY TEST, NUMBER 6

1. burn
2. lot
3. sub
4. home
5. dime
6. which (or witch)
7. keen
8. yes
9. boat
10. sure
11. hurl
12. door
13. kite
14. sell
15. nag
16. take
17. fall
18. week
19. death
20. love
21. tough
22. gap
23. moon
24. choice
25. king
26. size
27. pool
28. vine
29. chalk
30. laud
31. goose
32. shout
33. fat
34. puff
35. jar
36. reach
37. rag
38. mode
39. trip
40. page
41. raid
42. raise
43. bean
44. hash
45. limb
46. third
47. jail
48. knock
49. whip
50. met

APPENDIX C

DISTINCTIVE FEATURE CLASSIFICATIONS

APPENDIX C

SONORITY FEATURES

I. vocalic/non-vocalic

acoustically - presence (vs. absence) of a sharply defined structure;

genetically - primary or only excitation at the glottis together with a free passage through the buccal tract.

II. consonantal/non-consonantal

acoustically - presence (vs. absence) of a characteristic lowering in frequency of the first formant, a lowering which results in a reduction of the overall intensity of the sound and/or of only certain frequency regions;

genetically - presence (vs. absence) of an obstruction in the buccal tract.

Vowels are vocalic and non-consonantal. Consonants are consonantal and non-vocalic. Liquids are vocalic and consonantal (with both free passage and obstruction in the buccal cavity and with the corresponding acoustic effect). Glides are non-vocalic and non-consonantal; they never participate in the oppositions grave/acute and compact/diffuse and the basic or only glide of a given language is a one-feature phoneme in opposition to a phonemic zero.

III. nasal/oral (properly speaking, nasalized/non-nasalized)

acoustically - presence (vs. absence) of the characteristic stationary nasal formant with a concomitant reduction in the intensity of the sound and an increased damping of certain oral formants;

genetically - mouth resonator supplemented by the nose cavity (vs. the exclusion of the nasal resonator).

IV. compact/diffuse

acoustically - concentration of energy in a relatively narrow, central region of the auditory spectrum (vs. a concentration of energy in a non-central region), with a concomitant increase (vs. decrease) of the total amount of energy and its spread in time;

genetically - forward-flanged vs. backward-flanged. The difference lies in the relation between the shape and volume of the resonance chamber in front of the narrowest stricture and behind this structure. The resonator of the forward-flanged phonemes (wide vowels, and velar or palatal, including post-alveolar, consonants) is horn-shaped, whereas the backward-flanged phonemes (narrow vowels, and labial or dental, including alveolar consonants) have a cavity that approximates a Helmholtz resonator.

In vowel systems this feature often appears to be split into two autonomous features - compact/non-compact (higher vs. lower concentration of energy in the central region), and diffuse/non-diffuse (higher vs. lower concentration of energy in a non-central region).

V. abrupt/continuant

acoustically - silence (at least in the frequency range above the vocal cord vibration) followed and/or preceded by a spread of energy over a wide frequency region, either as a burst or as a rapid transition of vowel formants (vs. absence of abrupt transition between sound and "silence");

genetically - rapid turning on or off of source either through that swift closure and/or opening of the buccal tract which distinguishes plosives from constrictives, or through one or more taps which differentiate the abrupt liquids like a flap or trill / r / from continuant liquids like the lateral / l /.

VI. strident/non-strident (mellow)

acoustically - presence (vs. absence) of a higher intensity noise accompanied by a characteristic amplification of the higher frequencies and weakening of the lower formants;

genetically - rough-edged vs. smooth-edged: supplementary obstruction creating edge effects (Schneidenton) at the point of articulation distinguishes the production of the rough-edged phonemes from the less complex impediment in their smooth-edged counterparts.

VII. checked/unchecked

acoustically - higher rate of discharge of energy within a reduced interval of time (vs. lower rate of discharge within a longer interval), with a lower (vs. higher) damping;

genetically - reduced (vs. non-reduced) portion of air due to the stoppage of egressive as well as ingressive pulmonic participation. Checked phonemes are implemented in three different ways--as ejective (glottalized consonants, as implosives or clicks).

VIII. voice/voiceless

acoustically - presence (vs. absence) of periodic low frequency excitation;

genetically - periodic vibrations of the vocal cords (vs. lack of such vibrations).

PROTENSITY FEATURES

IX. tense/lax

acoustically - longer (vs. reduced) duration of the steady state portion of the sound, and its sharper defined resonance region in the spectrum;

genetically - a deliberate (vs. rapid) execution of the required gesture resulting in a lastingly stationary articulation; greater deformation of the buccal tract from its neutral, central position; heightened air pressure.

The role of muscular strain, affecting the tongue, the walls of the buccal tract, and the glottis, requires further investigation.

The difference between tense and lax phonemes parallels that between notes played legato and staccato, respectively.

TONALITY FEATURES

X. grave/acute

acoustically - predominance of the low (vs. high) part of the spectrum;

genetically - peripheral vs. medial: peripheral phonemes (velar and labial) have an ampler and less compartmented resonator than the corresponding medial phonemes (palatal and dental).

In the nasal consonants this feature is sometimes split into two autonomous features--grave/non-grave, and acute/non-acute--based on the interplay of the nasal murmur and oral release. The pitch of the resonator murmur effected in the nasal cavity plus the adjacent portion of the buccal cavity from the velic to the oral stricture is lower when the occlusion is made in the anterior part of the mouth cavity as compared to the structure in its posterior part. In / m / the two-fold low pitch is grave, in /   / acute, whereas in dental and velar nasals this opposition may be neutralized by the discrepancy between the gravity and acuteness of the two pitches (murmur and release or vice versa).

XI. flat/non-flat

acoustically - flat phonemes are opposed to their non-flat counterparts by a downward shift and/or weakening of some of their upper frequency components;

genetically - the former (narrowed-slit) phonemes, in contradistinction to the latter (wider-slit) phonemes are produced with a decreased back or front orifice of the mouth resonator and a concomitant velarization which expands the mouth resonator.

XII. sharp/non-sharp

acoustically - sharp phonemes are opposed to their non-sharp counterparts by an upward shift and/or strengthening of their upper frequency components;

genetically - the former (widened-slit) phonemes, are produced with a dilated back orifice (pharyngeal pass) of the mouth resonator and a concomitant palatalization which restricts and compartments the mouth cavity.