THE EFFECTS OF INTENSITY UPON THE INTELLIGIBILITY OF TIME-COMPRESSED CNC MONOSYLLABLES

THESIS FOR THE DEGREE OF M. A.
MICHIGAN STATE UNIVERSITY

SHELLEY SCHWIMMER
1 9 7 1

22 — 012

d — 038

K — 012

n — 277

ABSTRACT

THE EFFECTS OF INTENSITY UPON THE INTELLIGIBILITY
OF TIME-COMPRESSED CNC
MONOSYLLABLES

By

Shelley Schwimmer

A review of the literature suggests potentially valuable theoretical and diagnostic implications for the use of time-compressed speech stimuli with neurological auditory disorders. The few reports to date, however, have failed to control for standardization of stimuli and/or have utilized questionable compression procedures and experimental methodology. Such variables as sensation level, right and left ear effects, and stimulus reliability, as measured using standardized stimulus items, have received little attention in this area.

The purpose of this study was to investigate the response accuracy of normal listeners to time-compressed (TC) speech presented at various sensation levels, using the four lists of Form B of the Northwestern University Auditory Test No. 6 (NU-6).

Each of the four lists of 50 monosyllables was time compressed by 30% through 70%, in 10% steps. Also, there was a 0% TC control condition. The compression was performed using the Zemlin-modified version of the Fairbanks Time Compressor. This resulted in 24 experimental tapes, all of which were constructed under highly controlled experimental conditions.

THESIS

Ninety-six right-handed normal hearing young adults, six groups of 16 members each, participated as listeners in this study. The members of each group received lists one through four of the NU-6 test under a specified TC condition. Each list was presented to the listeners at one of the following sensation levels: 8 dB, 16 dB, 24 dB, or 32 dB. The lists were counter-balanced relative to the sensation levels, for each 16 member group. In addition, the stimuli were randomly presented to eight of the sixteen members of each group via the right ear, and eight via the left ear. All listening procedures were consonant with strict experimental and clinical criteria.

The results revealed that as the TC percentage increased, listener accuracy decreased. This effect was most pronounced at the sensation levels of 8 dB and 16 dB. Further, the spread of average articulation scores among the first five conditions of time compression decreased at the higher sensation levels. The most dramatic decline in listener accuracy occurred at the highest time compression ratio for all sensation levels. At each time compression ratio, the greatest amount of list variability occurred at 8 dB SL, with generally decreasing variability at the higher sensation levels. Likewise, the greater the TC ratio, the greater the associated list variability. In addition, there appeared to be a trading relationship between sensation level and TC ratio, such that listener inaccuracy due to the distortion of one factor could be offset by improving the level of the other factor. Further analysis revealed no effects due to ear preference. In view of past research, this finding was not unexpected considering the fact that this task utilized monosyllabic stimuli under monotic listening conditions.

The results of this study provided needed information relative to TC speech and its theoretical and clinical significance. First, the results suggest that the Zemlin-modified version of the Fairbanks Compressor provides a more distortion-free signal than its predecessor. Secondly, the results suggest relationships between TC ratio and sensation level, which to date have not been elucidated, especially under controlled experimental conditions. These relationships provide the basis for theoretical speculation. In addition, the results provide normative data as a basis for future investigations of this measurement procedure as an aid to audiological evaluation of individuals with neurologically-based disorders of the higher auditory pathways. These and other theoretical implications are discussed.

THE EFFECTS OF INTENSITY UPON THE INTELLIGIBILITY OF TIME-COMPRESSED CNC MONOSYLLABLES

By

Shelley Schwimmer

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

1971

677190

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 Master of Arts degree.

Director of Thesis

Guidance Committee: Chairman

ACKNOWLEDGMENTS

I wish to extend my gratitude to Dr. Daniel S. Beasley, my thesis advisor, and to Dr. William F. Rintelmann and Dr. Herbert J. Oyer, the members of my guidance committee, for their valuable assistance in the preparation of this thesis.

I would also like to express my appreciation to Dr. Willard R. Zemlin, for his time and effort spent in processing the experimental tapes, and to Miss Lisa Holstead, for her hospitality during this time.

I further wish to thank all of my subjects who were kind enough to volunteer much of their own valuable time in order to participate in this study. Special thanks are extended to Barbara Maier and Barbara Mackey, for their constant encouragement throughout this endeavor.

Finally, I would like to express my sincere gratitude to Dr. Beasley, whose role in the creation of both this student and this study was much greater than he could ever hope to realize.

This study was supported in part by the Rehabilitation Services Administration--Traineeship Grant No. 526-T-71

TABLE OF CONTENTS

TABLE OF CONTENTS--continued

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Differential audiometric techniques presently used in clinical diagnosis have been proven useful in the detection of lesions in the peripheral auditory system and in the eighth cranial nerve. The standard battery of clinical tests (including pure tone, speech audiometry, loudness balancing, SISI, and Bekesy audiometry) yield response patterns which often are peculiarly associated with damage to a particular anatomical site.

Lesions in the brain-stem and the auditory cortex, by contrast, elude detection by existing audiometric procedures. Present evidence suggests that individuals with disorders of the higher auditory pathways usually exhibit response behavior which appears essentially normal (Willeford, 1969; Katz, 1969). It has become apparent that conventional auditory tests lack the structure and sensitivity necessary for the identification of lesions in the higher auditory centers.

In the diagnosis of central auditory dysfunction, investigators have abandoned traditional audiometric techniques in favor of specially designed measures which require cortical integration of complex signals in order for the subject to respond appropriately (Willeford, 1969). The function of the higher auditory centers, according to Bocca and Calearo (1963), is to organize simultaneous or successive elements of the acoustic speech signal into a definite pattern. It is recognized that the central auditory pathways provide a sufficient degree of intrinsic redundancy, even when damaged, to allow simple psycho-acoustic elements, such as pure tones, to satisfactorily reach the cortical integrative and interpretive centers (Bocca and Calearo, 1963). Miller and Licklider (1950), using distorted speech signals, demonstrated that there existed an excess of auditory cues needed for adequate intelligibility. They used a switching arrangement whereby the signals were electronically switched on and off. It was found that as long as the interruptions occurred more than ten times per second, the interrupted speech signal was easily understood. The intelligibility of monosyllables did not drop below 90% until 50% of the speech signal was discarded.

Attention has therefore become centered on tests involving verbal stimuli which have been altered to render comprehension more difficult. Since the individual with a lesion in the auditory cortex experiences little difficulty with standard speech audiometry (Katz, 1969), it is probable that he utilizes the normally redundant information contained in the speech signal to compensate for his pathological deficit. A common goal in the formulation of test material to adequately diagnose central auditory lesions has therefore been to reduce the extrinsic redundancy contained in the auditory message so that it can no longer be used to compensate for the deficit in intrinsic redundancy in the centrally damaged auditory system (Willeford, 1969; Calearo and Lazzaroni, 1957).

Distorted Speech Tests

In order to increase diagnostic precision of patients with central auditory disorders, investigators have attempted to acquire differential responses to distorted speech signals (Jerger, 1964; Bocca, 1967; deQuiros, 1964; Bocca and Calearo, 1963; Calearo and Lazzaroni, 1957). It was theorized that when a message with reduced redundancy, such as distorted speech, is passed into a damaged pathway, intelligibility will suffer when compared to normal responses (Bocca and Calearo, 1963). Several different forms of distorted speech signals have been used. Katz (1962) used staggered spondaic word tests to evaluate central nervous system (CNS) disorders. Jerger (1964), Bocca et al. (1955), Calearo (1957), and Bocca (1955) experimented with filtered speech and reportedly were able to pinpoint temporal lobe lesions contralateral to the tested ear. Calearo and DiMitri (1958) and Calearo et al. (1962) used periodically switched speech and periodically switched noise and found that intelligibility increased as the signal-to-noise ratio decreased.

Another form of distorted speech used to assess auditory problems of the central nervous system has been time-compressed speech. Fournier (1954) stressed the importance of the time factor in speech discrimination and related the increased 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 when words are presented to the elderly at a high average syllabic rate there typically results an increased difficulty in intelligibility.

Finzi (1955) showed in a series of systematic tests with accelerated speech that in aged presbycusic subjects, the speech reception threshold, which is intensity at which fifty percent of the words can be accurately discriminated, may be reached only rarely, and that the threshold shift is much increased if compared to one of the normal subjects.

Using the above two findings as a basis, 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" (their phrasing) as stimulus material. These sentences were recorded at the rates of 140, 250 and 350 words per minute (wpm) to generate articulation curves for both groups of subjects. The average threshold shift for the normative subject group was between 5 and 10 dB at the 250 wpm presentation, and 10 and 15 dB for the 350 wpm presentation, but the essential shape of the articulation curve remained unchanged. The three curves obtained at progressively increased syllabic rates ran parallel with each other, demonstrating 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. It was further reported that when the accelerated sentences were presented to a group of subjects with temporal lobe lesions, these subjects yielded poorer articulation curves when the speeded message was sent to the ear contralateral to the lesion.

Although Calearo and Lazzaroni demonstrated that cortical damage may result in poorer discrimination scores, they also found the lengthening of synaptic time in the central auditory pathways, originally attributed to presbycusics by Fournier, to be a significant factor in impairing the ability to respond to time-accelerated speech. It should be noted that in trying to assess the conclusions drawn from this study, difficulties arise from the fact that the authors did not precisely describe the speech stimuli used in their study, nor did they explain the method of time-compression used to process their speech stimuli. These omissions necessarily preclude replication of the study which would attest to the validity and reliability of its conclusions.

deQuiros (1964) used accelerated speech material to test 20 normal subjects, 15 subjects with peripheral hearing losses, seven presbycusics, and several well-defined groups of adults and children with central disorders, including six subjects with retrocochlear lesions and 28 subjects with central lesions. 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. Each sentence was presented at the rates of 140, 250 and 350 wpm, using the same method of speech acceleration as did Calearo and Lazzaroni (1957), which lacks further description. 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. deQuiros found that for his normal subjects, there was usually a common curve for all three speeds of sentence presentation. In the cases where shifts were noted between speeds, they were within the range of 10 dB, with the exception of the threshold of detectability, which tended to remain constant. For his subjects with conductive hearing losses, the results obtained were either similar to those of the normal hearing subjects, or the entire family of curves was shifted to a higher intensity. In cochlear losses, normal results were sometimes obtained, but more often, the maximum articulation score for the 350 wpm condition was lower than those obtained for normal subjects. In presbycusic subjects, there was usually a shift in the threshold of detectability for all rates of sentence presentation.

His discussion of 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. Unfortunately, in view of the small number of subjects considered in each category of CNS disorders (e.g., brain-stem lesions, cerebellar lesions, central vestibular disturbances, etc.), 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.

Due to 1) the inadequate description of the process utilized for speech compression, and 2) the lack of standardization of speech stimuli employed in the above studies, their replication using a larger number of subjects under controlled conditions is impossible. Thus, if studies of this nature are to assume any clinical significance, it is imperative that they use controlled time compression procedures as well as clinically standardized test materials.

Controlled Time Compression Procedures

A controlled electromechanical procedure for time compressing speech stimuli was developed by Fairbanks, Everitt and Jaeger (1954). Essentially, the process entails 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 is 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, however, 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, contac' with th ape loop. Each head reproduces, as recordc , the ma al on tha portion of the tape with which it makes c act. When t cylinder is so positioned that one head is just losing ontact with the tape while the preceeding he s just m g contact with the tape, the segment of t tape that wrapped around the cylinder between the e two head ever makes contact with a reproducing head, and is t refore not reproduced. This is referred to as the interval of discard, and was found to be maximally effective when it was 15 to 20 msec 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 (I_a) is composed of the discard interval (I_d) and the recorded interval (I_r), such that

$$I_a = I_d + I_r$$

The sampling frequency, F_s, i.e., the rate at which the input signal is sampled, is

$$F_s = 1/I_a$$

The compression ratio, R_c, is then defined as

$$R_c = I_d/I_a = I_d F_s$$

Both R_c and I_d 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. (This procedure has been modified by Zemlin (1971), in order to eliminate low-frequency buzz at high compression ratios in the original Fairbanks apparatus, as well as to streamline the procedures involved).

Using this electromechanical apparatus, Fairbanks et al. (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. These messages were read at approximately 141 wpm, recorded and compressed electromechanically in time 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 subscoring results 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 affects comprehension of factual material.

Shriner, Beasley and Zemlin (1969) examined the identification and reaction time to frequency-divided, frequency-divided/time-restored, and frequency-divided/time-compressed speech of 50 elementary school children, using stimulus items taken from a clinically standardized discrimination test for children (Picture Identification for Children: A Standardized Index (PICSI), Nosca, 1965). The results of reaction time measured in the Shriner, et al. study demonstrated that increases in reaction time are accompanied by a corresponding decrease in the percentage of responses that are correctly identified. Although the overall mean reaction times were not found to be significantly different, significant differences were shown to occur when each of the distorted conditions were compared to the normal run. Thus, they concluded that when small differences in identification accuracy scores occur, the inclusion of accompanying reaction time measures does not appear to resolve these differences. If larger differences occur, as they did when each of the experimental conditions was compared to the normal run, then reaction time appears to be highly dependent upon percentage of words correctly identified.

The consideration of reaction time as it affects perceptual processing of auditory stimuli can be related to the temporally-based perceptual processing concepts, particularly as discussed by Aaronson (1967) and Beasley (1970). Aaronson speculated that the interval of silence (interstimulus interval) between two stimuli may be utilized in the decoding of individual perceptual units, and further, that the relationship between stimulus duration and the size of the interstimulus interval may be critical in the development of perceptual strategies. The use of sentential material by Fairbanks et al. (1957) in their investigation of time-compressed speech, fails to allow for consideration of the individual effects of interstimulus interval and stimulus duration on the comprehension of factual material.

According to Miron and Brown (1968) it is necessary to take into account the presentation rate of the stimuli in terms of syllables and words and the associated variability of these measurements, the distribution of phonation time, and finally, information measures of method content as stimulus parameters in speech compression. The interaction of all of these factors produce experimental variables which, to date, have been uncontrolled in the studies of the effects of time compression on the intelligibility of sentential material.

According to Miller (1951), accurate perception of a spoken word depends to a considerable extent upon the frame of expectations in which that word occurs. The use of speech, such as monosyllables, in place of sentences, would serve to reduce the amount of syntactic and semantic information available to the listener as cues for intelligibility of the stimulus. The listener's frame of expectations would necessarily have less of an influence in his perception of the words than it would if the words were arranged in a semantic and syntactically meaningful order. It should further be noted that studies of the time compression on intelligibility making use of sentential material, rather than word lists, necessarily confound the variables of interstimulus interval and stimulus duration (Beasley, 1970; Aaronson, 1967).

In essence, the effects of compressing either stimulus duration or interstimulus interval on the intelligibility of the speech stimulus cannot be determined with the sentence as the speech stimulus unit.

Controlled Clinical Stimuli

To isolate the effect of reduced stimulus duration on intelligibility, Fairbanks and Kodman (1957) presented Egan's phonetically balanced (PB-50) word list 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 influence on intelligibility when the stimulus words were presented at a comfortable listening level. However, this study cannot be considered of prognostic value as a clinical tool because of its psychoacoustic emphasis on the nature of the stimuli. Also, specially trained listeners were employed as the subject group rather than a clinical population of "naive" listeners. Further, it has been demonstrated that the PB-50 word list reflects low reliability and a wide range of word difficulty among the test items (Eldert and Davis, 1951). If time-compressed speech is to be used clinically, it is necessary to utilize more reliable and theoretically sound stimuli.

Although the W-22 word lists were devised (Hirsh et al, 1952) in an attempt to overcome the reliability problem, it was not until the development of the Northwestern Auditory Test No. 6 (NU-6) that the latter difficulty was investigated (Tillman and Carhart, 1966).

The development of this test was based upon studies performed with an earlier list, the Northwestern Auditory Test No. 4 (NU-4) developed by Tillman, Carhart and Wilbur (1963), which 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 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 as found 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 sensorineural 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 diagnostically differentiates among the various auditory pathologies.

Summary and Statement of the Problem

In summary, a review of the literature suggests theoretical and practical bases for the diagnostic value of time compressed speech stimuli in the identification of central auditory lesions (Bocca and Calearo, 1963; Calearo and Lazzaroni, 1957; deQuiros, 1964). However, several factors have served to undermine the clinical value of past investigations. First, there has been an inadequate description of the procedures used to compress the speech stimuli. Further, much of the research carried out with time compressed speech has made use of sentential material (Calearo and Lazzaroni, 1957; deQuiros, 1964; Fairbanks, et al. 1957) which, of necessity, involves the uncontrolled interaction of semantic and syntactic information as these

affect the intelligibility of the message. In addition, the effects of interstimulus interval and stimulus duration of the message *per se* could not be assessed independently with respect to their contribution to speech intelligibility. The replacement of sentences by speech stimuli consisting of isolated words obviates these variables. The use of a clinically standardized word list, such as that used in the clinical assessment of speech discrimination, would be of particular benefit in replicating past studies with different groups of subjects. Although the influence of stimulus duration on normative articulation scores has been examined using words from a standardized list (Fairbanks and Kodman, 1957), the effects of monosyllables has yet to be examined with respect to intensity variations in the stimulus presentation.

The purpose of this study, then, is 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 using normal hearing subjects. The evaluation of normative responses to words of compressed duration presented at various sensation levels will provide the groundwork upon which comparative studies with pathological subjects can be performed.

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

1) Will various degrees of time compression (30% to 70% in 10% steps) result in differential intelligibility scores?

2) Will the intelligibility scores at the various amounts of time compression interact with sensation level?

3). Will significant laterality effects result from varying the ear to which the stimuli are presented?

4) Are the results reliable, as measured by using equivalent versions of a standardized monosyllabic word list?

CHAPTER I

EXPERIMENTAL PROCEDURES

This study consisted of 96 subjects assigned to one of six experimental conditions. Each condition was characterized by one of six levels of time compression, utilizing four versions of a standardized word list presented at four sensation levels (SL), and according to a randomized right ear/left ear paradigm.

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. The test ear for each subject was determined by random selection of eight right ears and eight left ears for each of the six groups.

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 the normative status of his hearing bilaterally. In addition, a live-voice presentation of the CID W-1 Word List was administered unilaterally to obtain the Speech Reception Threshold for the designated test ear, i.e., the ear to receive the experimental stimuli.

Stimulus Generation

The experimental stimuli used in this study were the four lists of Form B of the NU-6 (Tillman and Carhart, 1966). Each list is composed of 50 CNC meaningful units, phonemically balanced according to the scheme advocated by Lehiste and Peterson (1962). 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).

A copy of each of the four recorded lists was made using 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). These copies were then temporally processed using the Fairbanks electromechanical time compression apparatus (Fairbanks, Everitt and Jaeger, 1954), as modified by Zemlin (1971). Each of the lists was time compressed by 30%, 40%, 50%, 60% and 70%. In addition, each list was passed through the time compression apparatus under the 0% time compression condition in order to control for the effects of possible fidelity distortion in the time compressed tapes. This resulted in 24 experimental tape recordings: six time compressed recordings for each of the four lists.

Copies of each experimental tape were made 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. During this procedure, approximately five seconds of silent interval response time was allotted between each stimulus item.

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 (see Appendix A), each list at one of four sensation levels: 8 dB, 16 dB, 24 dB and 32 dB. Within each group, the order of presentation of each of the four sensation levels was rotated. In this manner, each test list was presented at each sensation level a total of four times for each time compression condition, and the sensation levels were counterbalanced to avoid possible order effects of sensation level presentation.

All subjects were tested individually in a prefabricated double walled test chamber (IAC 1200 series) with the experimenter seated in the adjacent single-walled control room. The ambient noise in the test room was sufficiently low (45 dB on the C-scale of a Bruel and Kjar sound level meter) so as not to interfere with testing at even the lowest sensation level.

Each subject received standard instructions (see Appendix B) and was given a four page answer form (see Appendix C). The experimental tapes were then presented to the listener via a Viking Model 433 tape deck (frequency response 60-12,000 Hz, ±2 dB) mounted in a Maico MA 24 audiometric unit, through TDH 39 earphones mounted in MX 41/AR cushions.

Analysis

The data were hand-scored by the experimenter and converted to percentage correct scores. These scores were then plotted as articulation curves, and the slopes of these articulation functions were calculated for each condition of time compression. In addition, graphic data was computed for each of the other factors and respective interactions studied.

CHAPTER III

RESULTS

The results of this study generally support the thesis that as the ratio of time compression increases, the intelligibility of the word lists decreases. Further, as the sensation level increased, intelligibility was shown to increase under the several conditions of time compression. As for the inter-list equivalency of the NU-6, the results show that List I was the most difficult while List IV was the easiest. However, the effect of word lists interacted with sensation level and time compression. Finally, ear differences in this study were found to be minimal.

These results are reported in Tables 1 through 3 and depicted in Figures 1 through 5.

Time Compression and Sensation Level

Tables 1 and 2 and Figures 1 through 3 indicate that as the percentage of time compression increases, intelligibility decreases. An exception to this is the 0% and 30% conditions, whereby the 0% condition is slightly inferior at all sensation levels. 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 is a dramatic breakdown in intelligibility. This suggests that at high compression ratios, a normal listener may not have enough

Table 1. Average articulation scores for each condition of time compression at four sensation levels computed by test ear.

Percentage Time Compression

SL		0	30	40	50	60	70	$\bar{M}$ Total
8	L	60.2	67.0	48.8	51.2	50.2	17.0	49.0
	R	66.7	62.0	55.8	47.0	45.5	14.5	48.5
	M	63.4	64.5	52.3	49.1	47.8	15.7	48.8
16	L	81.3	84.8	79.0	71.2	68.8	31.5	69.4
	R	84.0	87.5	79.2	72.2	71.8	33.3	71.3
	M	82.6	86.1	79.1	71.7	70.3	32.4	70.3
24	L	89.5	91.8	88.2	88.5	81.2	49.5	81.4
	R	94.3	93.0	89.0	82.8	86.0	50.8	82.6
	M	91.9	92.4	88.6	85.6	83.6	50.1	82.0
32	L	95.8	96.8	94.8	91.2	90.8	69.2	89.7
	R	97.0	96.5	93.2	93.5	90.0	61.8	88.6
	M	96.4	96.6	94.0	92.3	90.4	65.5	89.2
$\bar{M}$ Total		83.5	84.9	78.5	74.6	73.0	40.9	72.5

time and/or information to perceptually process incoming verbal stimuli (Daniloff, Shriner and Zemlin, 1968; Shriner, Beasley and Zemlin, 1969; Aaronson, 1967).

Figure 1 and Table 1 also reveal that the intelligibility of time-compressed speech is significantly affected by sensation level. Specifically, listeners demonstrated increased discrimination ability at each condition of time compression as sensation level increased. This effect was most evident under 70% time compression, where the slope of the articulation curve is nearly linear. Under the other five percentages of time compression, the articulation functions are characterized by

Table 2. Average articulation scores for each condition of time compression at four sensation levels computed by test list.

Percentage Time Compression

SL	List	0	30	40	50	60	70	$\bar{M}$ Total
8	I	70.0	54.0	32.5	45.0	26.5	95.5	39.5
	II	66.5	77.5	54.5	50.0	47.0	14.0	51.5
	III	56.5	66.5	48.0	50.5	56.0	13.5	48.5
	IV	61.0	60.0	74.0	51.0	62.0	26.0	56.6
$\bar{M}$	Total	63.5	64.5	52.2	49.1	47.8	15.7	48.7
16	I	83.5	79.0	74.5	66.5	66.0	21.5	65.1
	II	82.0	84.5	76.0	69.0	57.0	32.5	66.8
	III	80.0	90.0	80.5	73.0	77.0	38.5	73.1
	IV	85.0	91.0	85.5	79.0	81.0	37.0	76.4
$\bar{M}$	Total	82.6	86.1	79.1	71.8	70.2	32.3	70.3
24	I	94.0	90.0	83.0	82.5	80.0	41.5	76.5
	II	92.0	94.0	90.0	86.0	84.5	43.0	81.5
	III	92.5	88.5	87.5	83.5	79.0	56.5	81.2
	IV	89.0	97.0	94.0	90.5	91.0	59.5	86.8
$\bar{M}$	Total	91.8	92.3	88.6	85.6	83.6	50.1	82.0
32	I	95.5	98.0	92.5	90.0	85.5	60.0	86.9
	II	97.0	95.5	93.5	95.5	91.5	64.0	89.5
	III	95.5	94.0	91.5	91.0	92.0	64.0	88.0
	IV	97.5	99.0	98.5	93.0	92.5	74.0	92.4
$\bar{M}$	Total	96.3	96.6	91.5	92.3	90.3	65.5	89.2
$\bar{M}$	Total	83.5	84.9	78.4	74.7	73.0	40.9	72.5

Figure 1.--Average articulation scores for six conditions of time compression plotted by sensation level

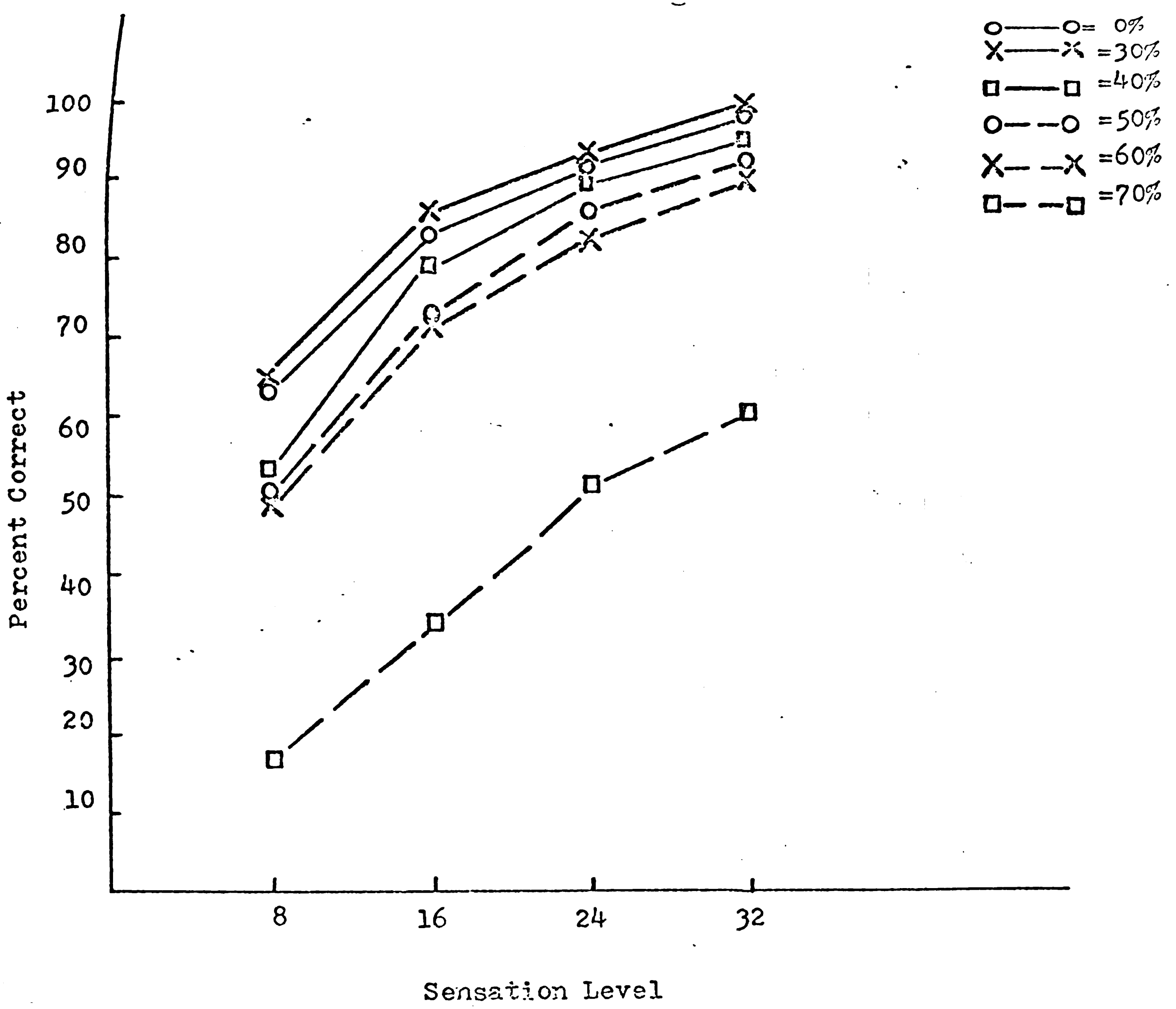
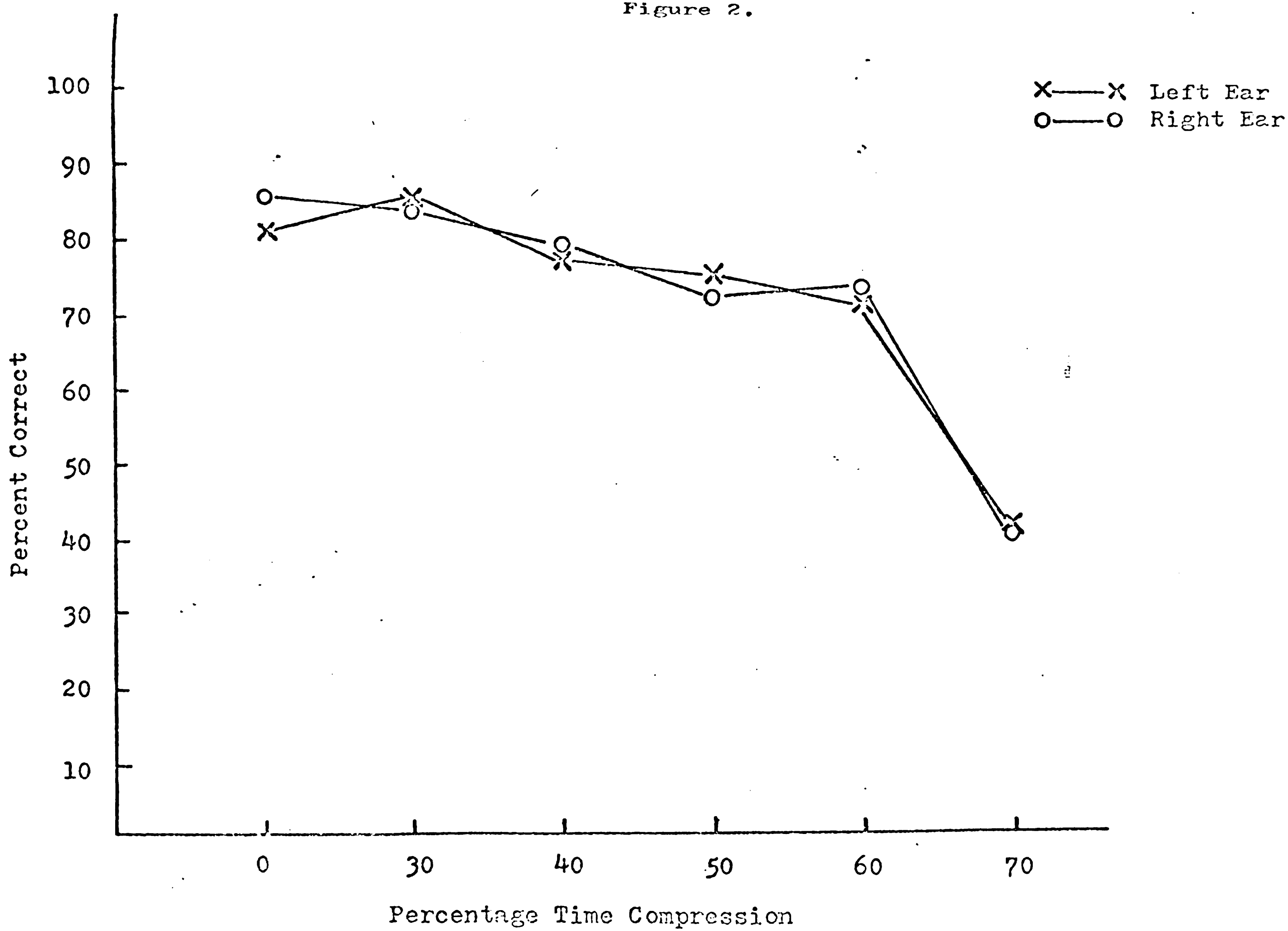

o——o= 0%
X——X =30%
□——□ =40%
o– –o =50%
X– –X =60%
□– –□ =70%
100
90
80
70
60
50
40
30
20
10
Percent Correct
8
16
24
32
Sensation Level

Figure 2.--Average articulation scores for right and left ear presentations plotted by time compression condition

Figure 2.


X——X Left Ear
O——O Right Ear
Percent Correct
100
90
80
70
60
50
40
30
20
10
0
30
40
50
60
70
Percentage Time Compression

igure 3.--Average articulation scores for right and left ear presentations plotted by sensation level

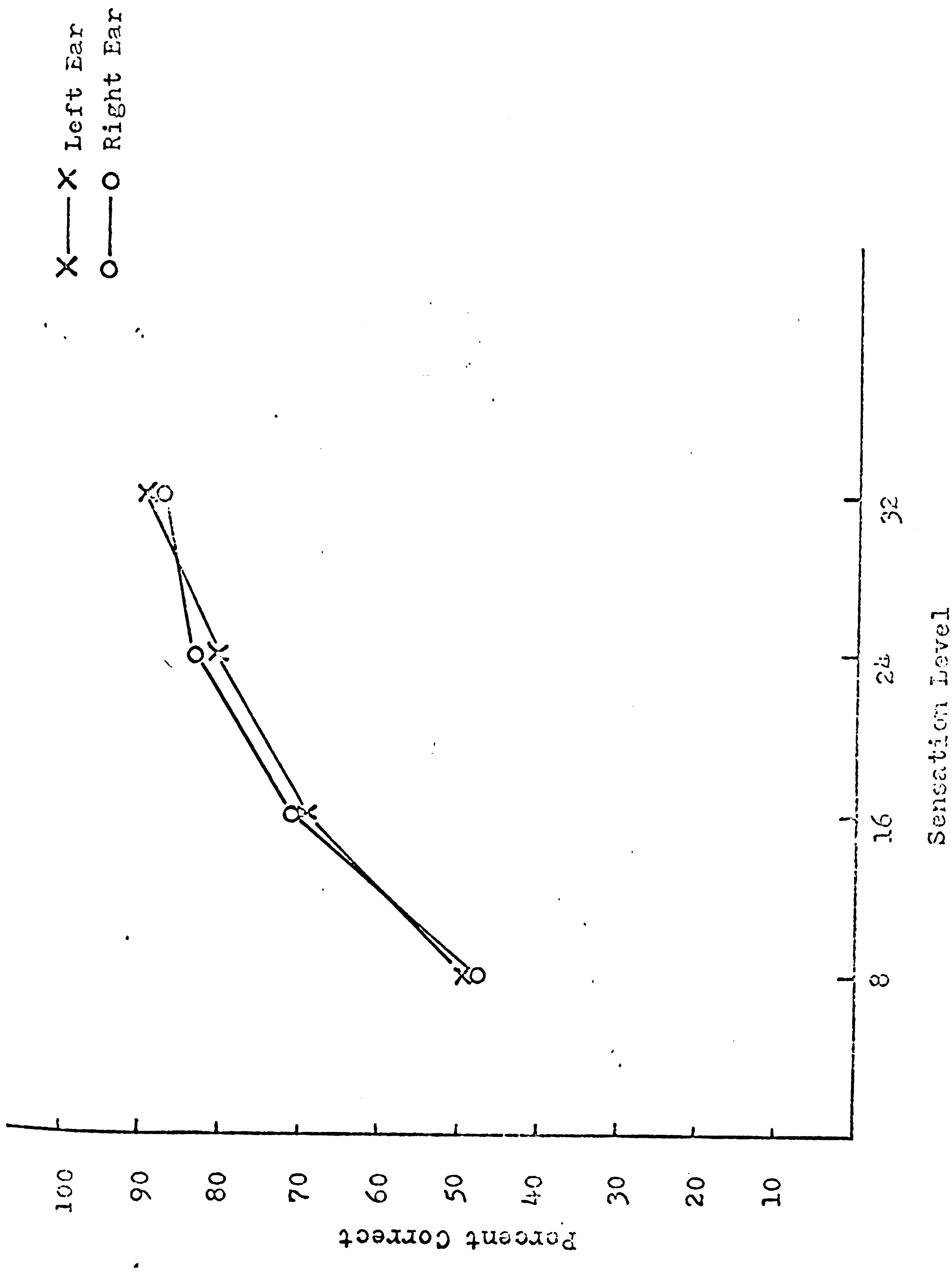

X——X Left Ear
O——O Right Ear
Percent Correct
100
90
80
70
60
50
40
30
20
10
8
16
24
32
Sensation Level

Figure 3.

rvilinear progressions in which discrimination scores crease less with the progressive elevation in intensity, aching an asymptote at 32 dB SL. The slopes of the articula-on curves generated for each condition of time compression re all found to increase between 2 and 3.5 % per dB increase intensity, when computed between the sensation levels of dB and 16 dB. This is consistent with previous data compiled undistorted NU-6 lists with respect to the slopes of the ticulation functions computed between these two sensation vels (Rintelmann and Schumaier, 1971). At the sensation vel of 32 dB, the mean discrimination scores under each ndition of time compression, with the exception of 70%, ll above 90% correct, which may be considered clinically rmal.

r Effects

Table 1 and Figures 2 and 3 depict the results according right and left test ear. These data suggest that there was sentially no difference between right and left ear presenta-on, at the several conditions of time compression and sensation vel under study. An interesting point, revealed in Figure 2, that the 0% time compression, left ear mean score appears to abnormally low. This could account, in part, for the anomaly scussed above, where it was noted that the 0% time compression ndition was slightly inferior to 30% time compression at the ur sensation levels.

List Effects

Tables 2 and 3 and Figures 4 and 5 illustrate the effects of the four lists of Form B of the NU-6 at the six conditions of time compression, the four sensation levels, and left and right test ear. Inspection of Figure 4 reveals a general decrease in intelligibility for all lists, except List IV, as the time compression ratio increases. Specifically, this effect occurs beyond 0%, whereas at 0% time compression, there is essentially no list differences (range from c. 81%-86%). The rather erratic configuration exhibited by List IV may be explained by simply noting that List IV will sustain a greater degree of distortion before intelligibility declines. Support for this contention is evident from Figure 4 and Table 2 where it can be found that List IV is the easiest, while List I appears to be the most difficult. However, because of the small spread of scores between lists overall, it can be concluded that the interlist differences are probably clinically negligible.

Reference to Table 2 and Figure 4 reveal that the interlist variability is greatest for the time compression ratios of 40%, 60% and 70%, whereas 30% and 50% show similar range of scores.

Reference to Table 2 and Figure 5 reveal that as sensation level decreased, the interlist variability increased. Again, List IV and List I appear to be the easiest and most difficult, respectively.

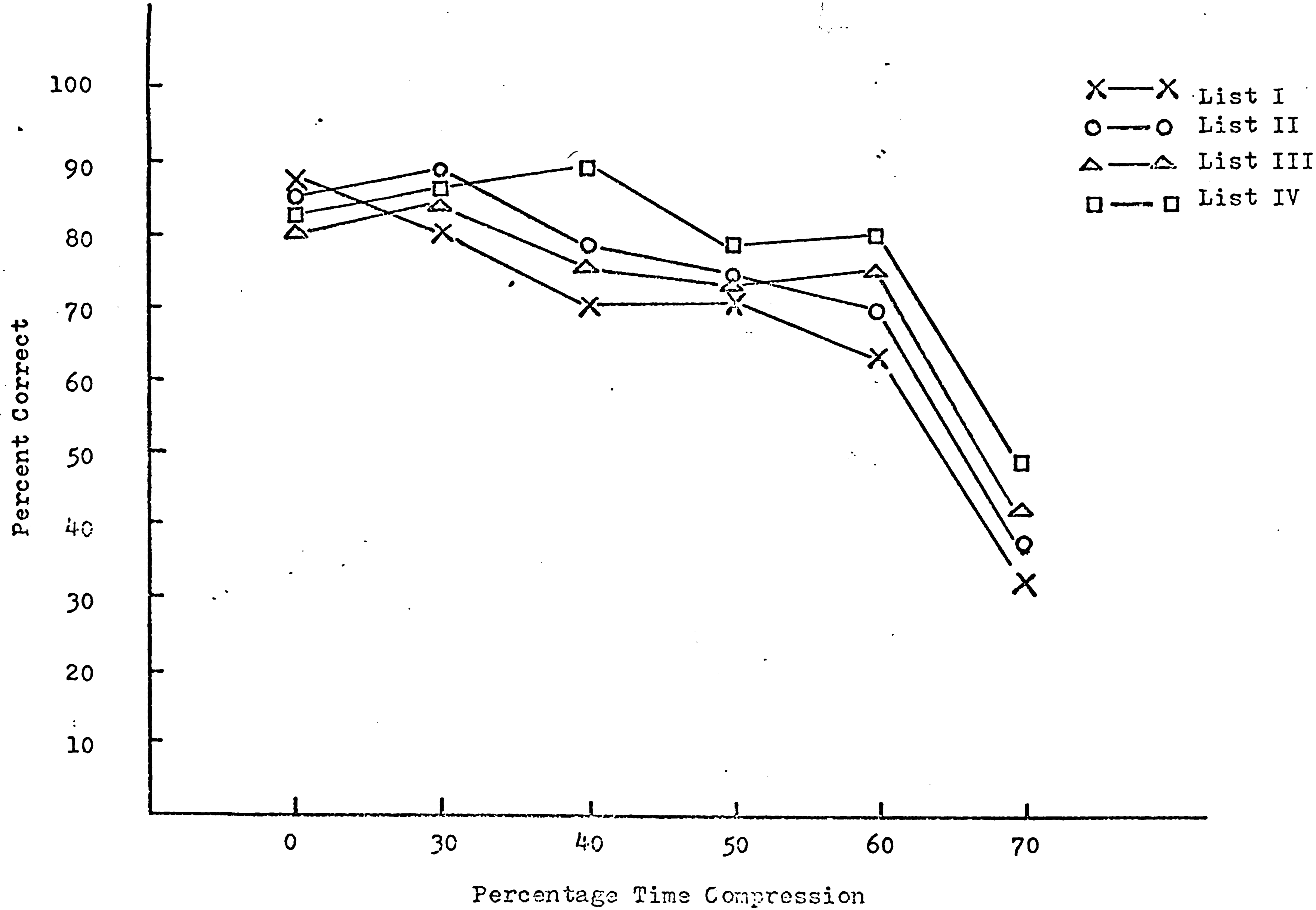
Table 3. Average articulation scores for right and left ears for four test lists at four sensation levels.

LIST

SL		I	II	III	IV	$\bar{M}$ Total
8	L	31.0	56.3	46.8	60.3	48.6
	R	47.0	46.7	69.3	50.2	53.3
	$\bar{M}$	39.0	51.5	58.0	55.2	50.9
16	L	60.8	60.0	76.2	73.3	67.5
	R	62.2	72.6	70.2	79.5	71.1
	$\bar{M}$	61.5	66.3	73.2	76.4	69.3
24	L	75.0	83.7	78.2	88.3	81.3
	R	81.8	79.1	83.8	85.3	82.5
	$\bar{M}$	78.4	81.4	81.0	86.8	81.9
32	L	91.2	87.3	89.5	91.1	89.7
	R	82.7	91.7	86.2	93.5	88.5
	$\bar{M}$	86.9	89.5	87.8	92.3	89.1
$\bar{M}$ Total		66.4	72.1	75.0	77.6	72.9

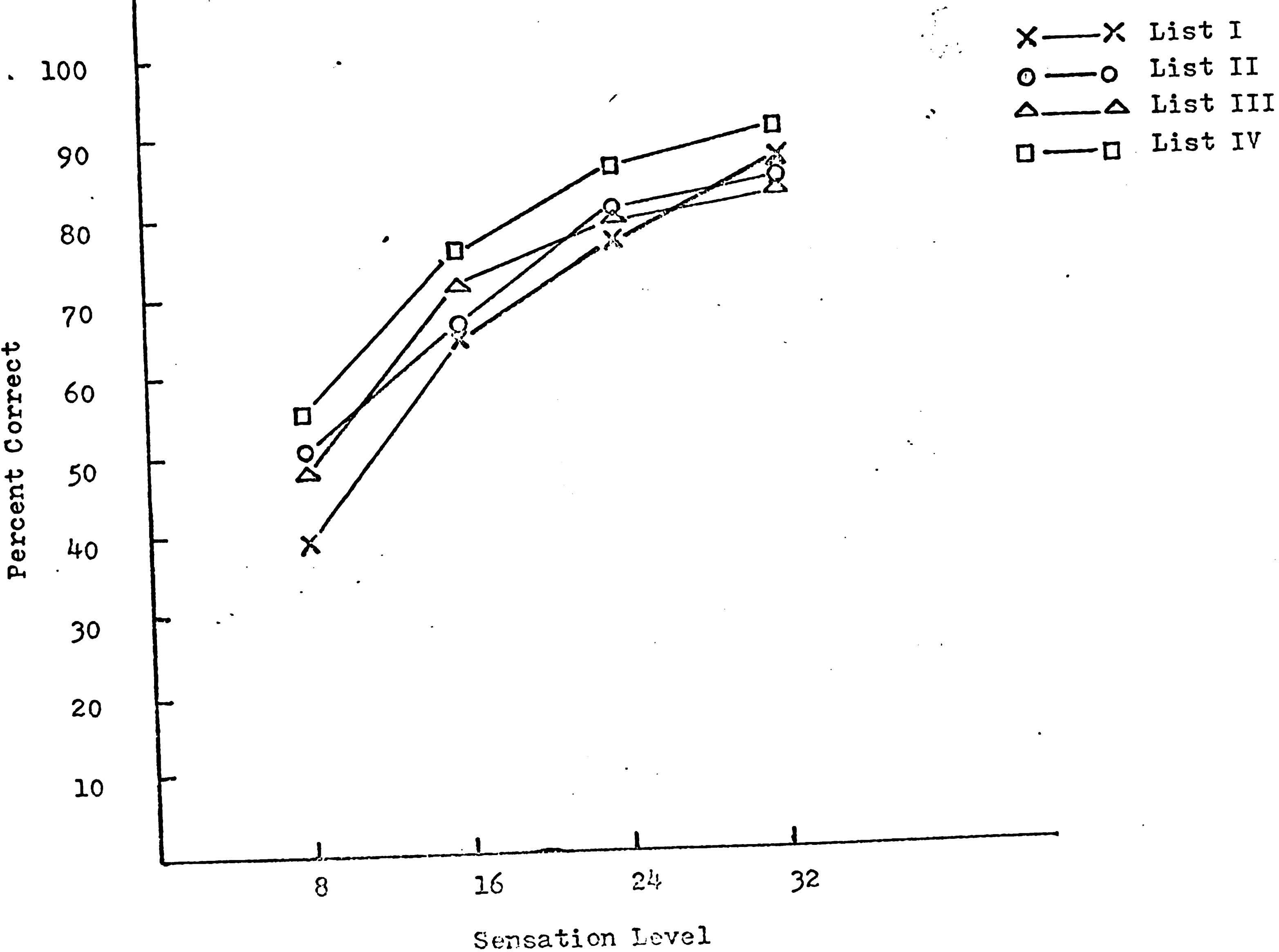
Finally, Table 3 indicates that there was essentially no consistent interaction between lists and test ear. However, this conclusion is highly tentative, due to the small and unequal number of subjects per cell which resulted from the fact that test ear was randomly assigned within each subject group.

Figure 4.--Average articulation scores for each of four test lists plotted by time compression condition



List I
List II
List III
List IV
100
90
80
70
60
50
40
30
20
10
Percent Correct
0
30
40
50
60
70
Percentage Time Compression

Figure 5.


X——X List I
O——O List II
△——△ List III
□——□ List IV
Percent Correct
100
90
80
70
60
50
40
30
20
10
8
16
24
32
Sensation Level

CHAPTER IV

DISCUSSION

Clinical Implications

The function of the auditory mechanism becomes progressively less definitive when consideration is given to the pathway central to the cochlear apparatus (Bocca and Calearo, 1963). According to Jerger (1960), the higher in the central nervous system the neurological lesion, the more subtle or complex are the stimuli required to uncover the lesion. In an analogy comparing the auditory system to a "bottleneck", he states that once conventional auditory tests using pure-tone and speech stimuli pass through the "bottleneck" of the peripheral auditory system up to and including the eighth ierve, more difficult material is necessary to determine the existence of neurological auditory disorders. Bocca and alearo (1963) and Bocca (1967) have suggested the value of listorted speech tests to identify the more subtle effects f lesions in the higher auditory pathways.

As one form of distorted speech signal, the potential linical utility of time-compressed speech has been demonstrated. alearo and Lazzaroni (1957), using time-compressed sentential aterial, found that with a group of subjects with ascertained ntrinsic lesions of the temporal lobe, the discrimination bility was clearly worse when the accelerated message was transmitted to the ear contralateral to the lesion. deQuiros (1964), using similar stimuli, found in various subject groups with central nervous system disorders, that accelerated speech testing provided useful information, which, when correlated with other findings, could aid in differential diagnosis of brain lesions.

An advantage of using temporally modified speech signals over other forms of speech distortion, such as filtering, in differential diagnosis of higher auditory lesions, is that the several theoretical models of speech perception suggest that perception is essentially temporally-biased (Aaronson, 1967). In addition, such signals do not eliminate possibly clinically relevant spectral information, such as formant structure, whereas this does not hold true for filtered speech.

Although there are several methods of temporally distorting speech signals, the most efficient and controlled method to date is time compression using the electromechanical compressor. This procedure has the additional advantage of having been utilized in a large number of studies (Fairbanks, et al, 1957; Fairbanks and Kodman, 1957; Daniloff, Shriner and Zemlin, 1968; Shriner, Beasley and Zemlin, 1969). If it is to be used clinically, however, it is necessary to obtain normative data using standard clinical procedures. This study has provided the necessary normative data. Before further investigations are carried out clinically, however, it is necessary to consider several pertinent points relative to the findings of this study.

Time Compression. The results of this study indicate a gradual decrease in the intelligibility of monosyllables corresponding to progressively greater percentages of time compression over the range of 30% to 60%, with a dramatic reduction of intelligibility occurring at the 70% time compression condition. These findings are in agreement with those of Daniloff, Shriner and Zemlin (1968), who found a significant breakdown in intelligibility to occur at 70% time compression using eleven different vowels placed in an /h-d/ context as their experimental stimuli. Fairbanks and Kodman (1957), however, found no appreciable breakdown in intelligibility of phonetically balanced monosyllables until a time compression ratio of 80% was reached. The discrepancy between the results of Fairbanks and Kodman (1957) and those of Daniloff, Shriner and Zemlin (1968) and those of the present study can be accounted for in several ways.

First, Fairbanks and Kodman (1957) used a ten msec discard interval in their method of time compression, as opposed to the 20 msec discard interval employed in the other two studies. According to Daniloff, Shriner and Zemlin (1968), this probably served to enhance intelligibility at the high compression ratios employed in these studies since smaller segments of the message and therefore, smaller bits of information, were deleted at a time. Secondly, Fairbanks and Kodman (1957) used highly trained listeners as subjects for their study, whereas "naive" listeners served as subjects in the other two studies. Finally, Fairbanks and Kodman (1957)

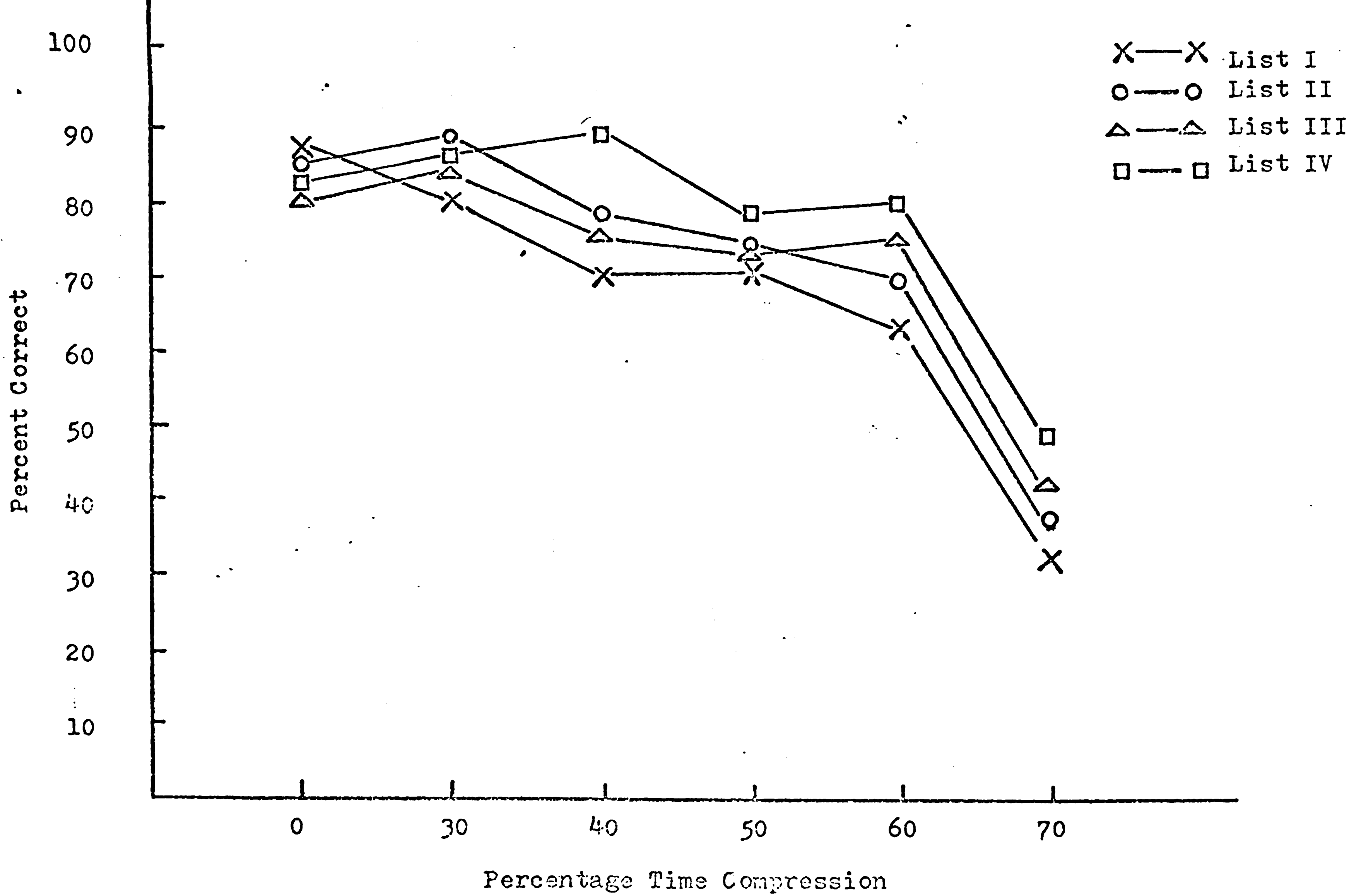
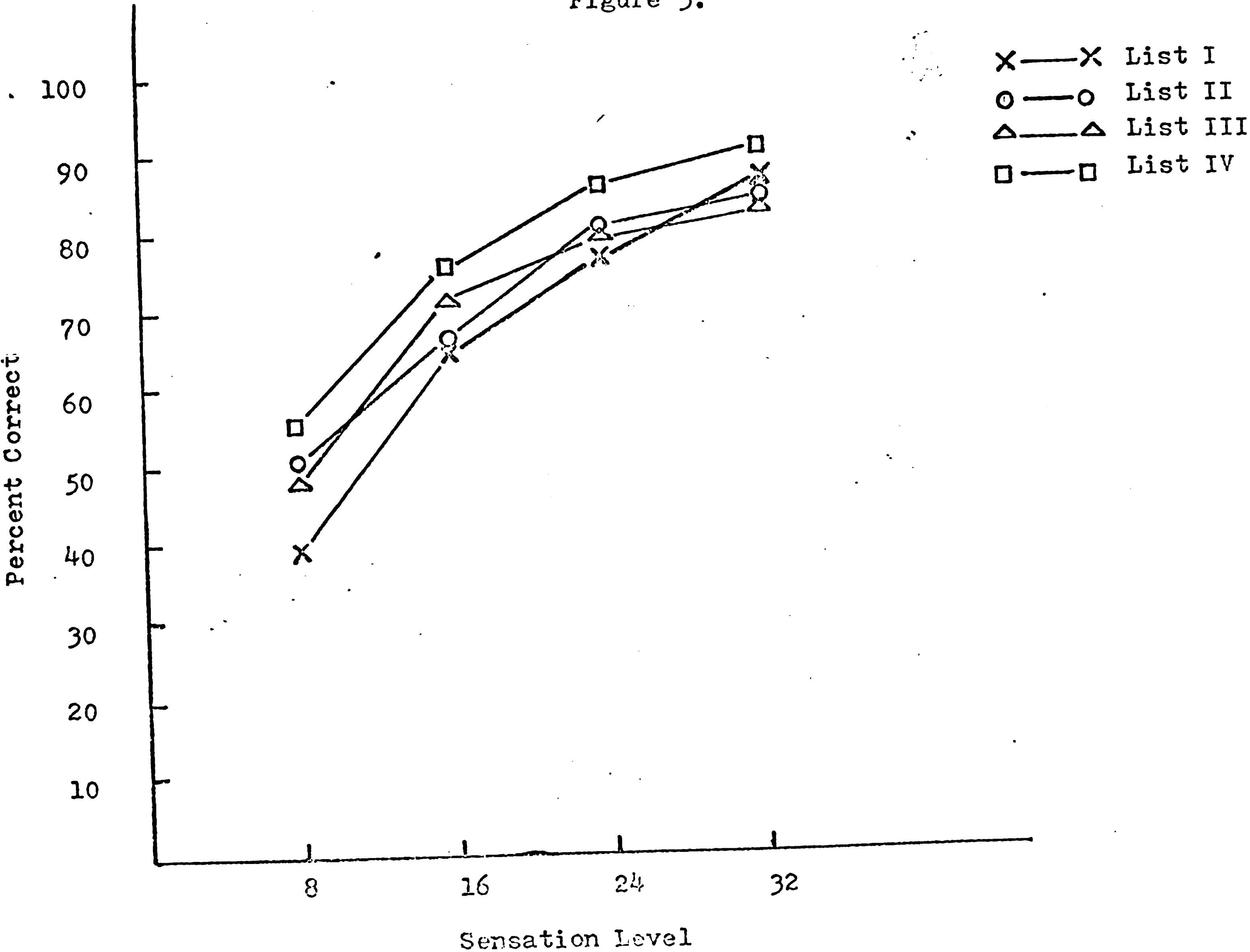

List I
List II
List III
List IV
100
90
80
70
60
50
40
30
20
10
Percent Correct
0
30
40
50
60
70
Percentage Time Compression

Figure 5.


List I
List II
List III
List IV
100
90
80
70
60
50
40
30
20
10
Percent Correct
8
16
24
32
Sensation Level

CHAPTER IV

DISCUSSION

Clinical Implications

The function of the auditory mechanism becomes progressively less definitive when consideration is given to the pathway central to the cochlear apparatus (Bocca and Calearo, 1963). According to Jerger (1960), the higher in the central nervous system the neurological lesion, the more subtle or complex are the stimuli required to uncover the lesion. In an analogy comparing the auditory system to a "bottleneck", he states that once conventional auditory tests using pure-tone and speech stimuli pass through the "bottleneck" of the peripheral auditory system up to and including the eighth nerve, more difficult material is necessary to determine the existence of neurological auditory disorders. Bocca and Calearo (1963) and Bocca (1967) have suggested the value of distorted speech tests to identify the more subtle effects of lesions in the higher auditory pathways.

As one form of distorted speech signal, the potential clinical utility of time-compressed speech has been demonstrated. Calearo and Lazzaroni (1957), using time-compressed sentential material, found that with a group of subjects with ascertained intrinsic lesions of the temporal lobe, the discrimination ability was clearly worse when the accelerated message was transmitted to the ear contralateral to the lesion. deQuiros (1964), using similar stimuli, found in various subject groups with central nervous system disorders, that accelerated speech testing provided useful information, which, when correlated with other findings, could aid in differential diagnosis of brain lesions.

An advantage of using temporally modified speech signals over other forms of speech distortion, such as filtering, in differential diagnosis of higher auditory lesions, is that the several theoretical models of speech perception suggest that perception is essentially temporally-biased (Aaronson, 1967). In addition, such signals do not eliminate possibly clinically relevant spectral information, such as formant structure, whereas this does not hold true for filtered speech.

Although there are several methods of temporally distorting speech signals, the most efficient and controlled method to date is time compression using the electromechanical compressor. This procedure has the additional advantage of having been utilized in a large number of studies (Fairbanks, et al, 1957; Fairbanks and Kodman, 1957; Daniloff, Shriner and Zemlin, 1968; Shriner, Beasley and Zemlin, 1969). If it is to be used clinically, however, it is necessary to obtain normative data using standard clinical procedures. This study has provided the necessary normative data. Before further investigations are carried out clinically, however, it is necessary to consider several pertinent points relative to the findings of this study.

Time Compression. The results of this study indicate a gradual decrease in the intelligibility of monosyllables corresponding to progressively greater percentages of time compression over the range of 30% to 60%, with a dramatic reduction of intelligibility occurring at the 70% time compression condition. These findings are in agreement with those of Daniloff, Shriner and Zemlin (1968), who found a significant breakdown in intelligibility to occur at 70% time compression using eleven different vowels placed in an /h-d/ context as their experimental stimuli. Fairbanks and Kodman (1957), however, found no appreciable breakdown in intelligibility of phonetically balanced monosyllables until a time compression ratio of 80% was reached. The discrepancy between the results of Fairbanks and Kodman (1957) and those of Daniloff, Shriner and Zemlin (1968) and those of the present study can be accounted for in several ways.

First, Fairbanks and Kodman (1957) used a ten msec discard interval in their method of time compression, as opposed to the 20 msec discard interval employed in the other two studies. According to Daniloff, Shriner and Zemlin (1968), this probably served to enhance intelligibility at the high compression ratios employed in these studies since smaller segments of the message and therefore, smaller bits of information, were deleted at a time. Secondly, Fairbanks and Kodman (1957) used highly trained listeners as subjects for their study, whereas "naive" listeners served as subjects in the other two studies. Finally, Fairbanks and Kodman (1957) presented their stimuli at sufficiently high sensation levels (80 dB SL) to insure maximum intelligibility, whereas for the purposes of the present study, the maximum intensity level employed was at a sensation level of 32 dB. At this level, PB Max was not reached for the 70% time compression condition.

An exception to the tendency for intelligibility to decline with gradual increments in the time compression ratio of monosyllables in this study seems to occur between the 0% and 30% time compression conditions. Equivocal as it is, this is not the only instance in which such a finding has occurred.

Aaronson (1966) conducted a study in which the duration of the auditory test stimuli was varied for a fixed presentation rate. Spoken digits were compressed by 35% and compared to the original seven digit sequence with respect to ease of short term recall. It was found that when subjects memorized each sequence and tried to recall it a few seconds after presentation, they did better with the sequence of digits that were time-compressed by 35% than they did under conditions of 0% time compression. Zemlin, Daniloff and Shriner (1968), in their psychological scaling of time compressed speech, found that normal listeners prefer speech at about 30% time compression, although it was scaled as being more difficult to understand.

These findings suggest the possibility that the amount of redundancy (or lack of it) present in the undistorted speech signal may inhibit, if only to a slight degree, the discrimination ability and/or the comprehension ability of normal listeners, and that an optimal listening duration of signals compressed between 30 and 35% may facilitate information processing. Future research is needed to clarify this issue.

Sensation Level. The interaction of intensity and the effects of time compression suggest a relationship between the two factors. The data obtained from the present study indicate that for every increase in intensity, there is a corresponding increase in intelligibility of the test stimuli for all conditions of time compression employed. The effects of intensity increments lessen, as would be expected, as an optimal listening intensity is approached. This effect is readily observable from the curvilinear configuration of the articulation functions of the first five levels of conditions of compression. (Figure 1).

The effects of intensity on the comprehension ability of normal subjects used in the studies of Calearo and Lazzaroni (1957) and of deQuiros (1964) cannot readily be compared to the present findings due to the difference in speech stimuli and time compression methods employed by each. It should be noted, however, that a tendency for intensity to neutralize the effects of speech acceleration was reported by Calearo and Lazzaroni (1957) for their normal subjects. This holds true to a considerable extent in the present study as well. Although the mean articulation scores for each of the first five conditions of time compression are not exactly equal even at the highest sensation level, they all fall within the clinical range of normal. This is not the case with the 70% time compression condition. At this percentage of time compression, the articulation function is still linear at the highest sensation level, indicative of the fact that sufficient intensity levels have not been reached for maximum discrimination. It is doubtful, in fact, that such a level would be reached for this degree of time compression.

Since tests of discrimination are clinically administered at sufficiently high intensity levels for the patient to perform optimally (Hirsh, 1952), it can be reasoned, on the basis of this study, that a clinical test of discrimination using time-compressed monosyllables, cannot employ stimuli compressed beyond 60% at the highest sensation level examined. To do so would allow for the interference of accurate determination of speech discrimination ability. Further studies using pathological subjects should yield valuable information relative to the precise combinations of time compression and intensity levels which would be of maximum diagnostic utility in the formulation of such a discrimination test.

Test Ear. In order to utilize validly the same test for both right and left ears, performance of normal subjects would warrant that test results between ears be approximately equal. If this were not, in fact, the underlying assumption of all audiometric testing, then separate tests would have been devised for each ear, taking into account laterality effects on the respective scores of each ear. The question of differences between ears is therefore worthy of consideration in light of the potential utility of time-distorted speech as a diagnostic tool for central auditory disorders. The absence of ear differences would allow for the utilization of a single test.

There is considerable evidence, however, that under dichotic and monotic listening conditions, certain kinds of material are recalled or recognized better if they are received by one ear rather than the other. It has further been found that the ear of preference is dependent upon the stimulus used.

Under dichotic listening conditions, digits (Kimura, 1961; Broadbent and Gregory, 1964), phonetically matched words (Borkowski, Spreen and Stutz, 1965) and plosive consonants (Shankweiler and Studdert-Kennedy, 1967) are better received in the right ear, whereas orchestrated melodies (Kimura, 1964), sonar signals (Chaney and Webster, 1966), environmental sounds (Curry, 1967), 2-click thresholds (Murphy and Venables, 1969) and rapid pitch changes (Darwin, 1969) are better received in the left ear.

If differential ear preference is a function of cerebral dominance, as Kimura (1961) contends, and it has been found to be independent of order in which the signals to the two ears are recalled (Bryden, 1963), then it is reasonable to assume that there are differences in the ability of the two hemispheres to process different types of sounds. Because it is assumed that each ear has greater neural representation in the opposite cerebral hemisphere, the predominance of the left hemisphere for speech, for example, is reflected in superior recognition for words arriving at the right ear (Kimura, 1967). In effect, the above studies (Kimura, 1961; Broadbent and Gregory, 1964; Borkowski, Spreen and Stutz, 1965; Shankweiler and Studdert-Kennedy, 1967) with dichotically presented pairs of speech stimuli support this contention.

Studies using monotically presented speech stimuli have investigated the question of whether the hemispheres, when stimulated unilaterally, will function comparably with each other, or whether, in fact, there is a dominance of one hemisphere for particular stimuli, as has been evidenced in dichotic listening tasks. If the latter holds true, then superior performance of the ear contralateral to the dominant hemisphere for that stimulus could be expected. Dirks (1964) presented filtered phonetically balanced words both dichotically and monotically to normal listeners. He obtained a significant right ear superiority under the dichotic listening condition, but only a very small, nonsignificant right ear superiority under the monotic condition. Kimura, using both words and nonsense syllables, found significant right ear advantages when subjects reported only one ear at a time in both dichotic listening tasks, but no significant difference in the monotic tasks. Corsi (1967) found no significant ear advantages in three different tests of speech perception when they were presented monotically.

It should be noted that in all of the above monotic studies, only Kimura's word study failed to find at least some small difference in favor of the right ear. It could thus be argued that there may, in fact, be a right ear superiority under monotic presentation conditions, but one which was too small to be detected in a statistical analysis

of the above studies. This contention is substantiated by the results of a survey of 3,465 subjects (Glorig, et al, 1958). A slightly lower right ear threshold for both speech and pure tones was found with monotic presentation of the stimuli, but this was of a lower magnitude than is typical under dichotic listening conditions. Thus, it is necessary to consider ear effects in clinical audiometry, although to date, such effects appear to be minimal in monotic listening tasks. Since this study also revealed no ear effects, it appears that it is clinically useful monotically and will not be confounded by ear laterality effects. This conclusion is tentative, however, due to the small and unequal number of subjects per right or left ear condition presented with each of the four test lists at a given sensation level in this study. Further research is thus indicated to establish its validity.

List Effects. All lists are essentially equal at 0% time compression whereas higher ratios of time compression as well as lower sensation levels serve to increase the interlist variability. Further, List IV overall is the easiest and List I is the hardest. It should be recognized, however, that the order of list presentation remained constant for all subjects under each condition of time compression. Despite the fact that intensity order was counterbalanced within each of these conditions, it is possible (see Figure 4) that order effects of presentation come into play under the conditions of greater degrees of time compression, establishing Lists I, II, III, and IV as hardest to easiest, respectively. Therefore, in the formulation of a clinical discrimination test using time-compressed speech stimuli, the effects of adapting to the discrimination task, resulting in improved scores is worthy of further investigation in order to avoid confusing the effects of practice with discrimination ability.

Theoretical Implications

Temporal considerations with regard to the processing of speech stimuli have received a considerable amount of attention in the literature, much of which has been focused on the interpretation of short-term memory functions. Theoretical models typically describe two stages of perception involved in the performance of immediate recall tasks (Pollack, 1959; Mackworth, 1959; Sperling, 1960, 1963; Sternberg, 1964; Broadbent, 1957a, 1957b, 1958).

In the first stage, the stimulus representation are thought to be unidentified sensations which are unstable and subject to change over the passage of time or as a result of intervening events. These representations are held in a large capacity perceptual storage system which can receive more than one item simultaneously. In the second stage, processing or encoding of the individual stimulus items taken from the storage system occurs. Each element is encoded one at a time and these representations are more permanent than those in stage one.

In most models of short-term memory, according to Aaronson (1967), the extent to which errors occur during perception rather than during subsequent retention or retrieval of the information cannot be determined directly. The roles of stimulus duration and the rate of stimulus presentation have been speculated upon in relation to the aforementioned theoretical framework. Aaronson (1967) for example, suggests that stimulus duration might determine the amount of stimulus information that goes into a buffer storage (stage one), while duration of the interstimulus interval might affect time available for encoding the representation. She further contends that the amount of time between stimulus items, during which the stimulus is available to the subject for perceptual processing, may be a more important factor than physical stimulus duration in determining recall accuracy. In support of this contention, she cites the study by Fairbanks, et al. (1957) in which subjects were presented time-compressed messages and messages of normal rate. The comprehension ability was judged to be poorer for the condition of time compression than that for the normal rate. To isolate the effects of stimulus duration on the intelligibility of the message, Fairbanks and Kodman (1957) used time-compressed monosyllables presented at uniform interstimulus intervals to normal listeners. They found that 80% of the original duration of the words could be discarded before intelligibility would drop below 95% correct.

The inference drawn by Aaronson from the two preceding studies concerning the relative importance of interstimulus interval over stimulus duration in the processing of auditory information is contraindicated to an extent by the results of the present study. The present findings reveal that once a

ritical duration is reached, at which point between 60% and
0% of the signal is deleted, the intelligibility of the
onosyllabic unit is significantly reduced. Further, the
iberal amount of time provided after each word presentation,
orresponding to what would be considered the encoding stage,
id not appear to compensate for the reduced informational
edundancy provided by the acoustic signal which was thus
educed in duration.

Beasley (1970) provided further indication of the
portance of word duration as opposed to interstimulus
terval duration in the perception of various orders of
ntential approximations. He states, ". . . increasing
e interstimulus interval does not necessarily offset the
gradating effects of a decrease in word duration in
ntential stimuli" (p. 36). His findings indicate, in
ct, that word duration was a more crucial variable in
call accuracy.

Thus, in tasks not involving comprehension, the rela-
ve importance of the interstimulus interval in neutralizing
e effects of reduced stimulus duration in minimized. In the
ocessing of complex material, the use of the interstimulus
terval to process semantic as well as non-verbal cues may be
considerable importance. In the intelligibility of isolated
ds or synthetic sentences, however, where semantic constraints
not as great, interstimulus interval duration cannot compen-
e for the reduced word duration in improving performance once
duration of the word has been reduced beyond a critical
ation.

Although the study by Beasley is not exactly comparable to the current study due to stimulus differences, it does indicate that the use of complex strings of verbal stimuli, such as synthetic sentences (Speaks and Jerger, 1965; Beasley (1970), as a stimulus for testing higher auditory centers, may be unnecessary. That is, the monosyllabic lists, which provide less variability than sentential material, may be used for this purpose if they are appropriately distorted. Further work on pathological populations should provide valuable information relative to this contention.

Implications for Further Research

Perhaps the most obvious area requiring further research is that concerned with the responses of a clinical population to the stimuli used in this study. Specifically, the effects of time-compressed monosyllables on the discrimination ability of subjects with conductive, cochlear, retrocochlear, and central nervous system disorders must be examined to assess its actual utility in the differential diagnosis of central auditory lesions. Further, such an investigation would serve to derive precise combinations of time compression and sensation level which would be considered most diagnostically valuable.

Another potential area of investigation lies in the comparison of time-compressed monosyllables with other forms of distorted speech signals developed for the assessment of central auditory lesions. A discussion of specific techniques is of little value at the present time since there are variations to each type of test. Further, standardization of stimulus material, procedures and instrumentation for these tests is presently lacking. On the basis of current research, however, each distinct method of speech distortion provides diagnostic information relative to particular sites in the central auditory system. To date, frequency distortion has been found to be most useful in the identification of lesions above the third order neuronal level, whereas interrupted, switched and time-distorted speech signals are more indicative of lesions at the second order neuron level (Bocca and Calearo, 1963). Additional diagnostic utility has been indicated for time-distorted speech, however, in the identification of diffuse damage of the higher auditory pathways and of lesions occurring at or above the level of the third order neuron, resulting in adverse test scores in the ear contralateral to the site of lesion (Calearo and Lazzaroni, 1957; deQuiros, 1964). The roles of each form of speech distortion can be viewed in the context of a test battery, for potential use in site of lesion testing. The attempt of this study to use standardized clinical methodology and stimulus items in the presentation of time-compressed speech stimuli can be seen as a springboard towards the development of such a battery.

The findings of this study suggest the need for investigating the effects of prolonged exposure to time-compressed speech stimuli on the discrimination ability of the listener.

hum
a d
indi
impr
horo
he c
ther
to i
exam
in i
if t
the
four
sub

res
the
the
con
wou
the
the
sup
nor
by
towa
the

Present results using four clinically standardized word lists found to be essentially equivalent in difficulty under normal conditions, suggest that under high percentages of time compression, the relative difficulty of the listening task decreased with each successive word list presentation. Since the order of list presentation was not rotated in this study, there remains a need to determine whether this effect is due to inequalities in the difficulty of the word lists under the examined conditions of time compression, or whether, in fact, it is due to order effects of presentation. It would appear, if the latter should be the case, that lengthy tests employing time-compressed stimuli should be avoided so as not to confound learning effects with an accurate appraisal of the subject's discrimination ability.

A further area of investigation indicated by the results of this study is related to the finding that, for the left ear, maximum articulation scores were obtained at the 30% time compression condition, rather than at the condition of 0% time compression, as would be expected. This would imply that perhaps the acoustic redundancy contained in the speech signal serves to impede the optimal functioning of the auditory system relative to its processing of speech. In support of this contention, prior research has indicated that normal hearing subjects preferred listening to speech compressed by 30% (Zemlin, Daniloff and Shriner, 1968). Research directed toward the question of whether an optimum signal duration for the auditory processing of speech actually exists must necessarily take into account the context of the speech signal. However, definitive findings concerning this issue may hold significant implications for the training of individuals with auditorily based disorders of language processing.

Finally, the matter of ear differences relative to the reception of time-compressed speech stimuli has only been incidentally investigated in the present study. Due to the random selection of test ear for each condition of time compression, the number of subjects administered each of the four lists at the same sensation level differed for each ear. In addition, the small number of subjects used to examine ear effects for each condition of time compression does not allow for a generalization concerning ear effects to be made at this time. More rigorous investigation of possible laterality effects, using larger and equal numbers of subjects for each test ear, is warranted in order for a definitive statement concerning laterality effects in the auditory processing of time-compressed speech stimuli to be made.

CHAPTER V

SUMMARY

The results of this study, within the limits of the experimental procedures employed, provide significant information relative to the intelligibility of time-compressed speech stimuli. It was found that increases in the compression ratio up to the level of 60% generally resulted in gradual decreases in the intelligibility of the word lists. An exception to this was found with the 0% and 30% conditions of time compression, in which case the mean left ear articulation scores for the 30% time compression condition exceeded the mean articulation scores for the 0% time compression condition at all sensation levels. Normative articulation functions, graphically depicting the effects of intensity upon the intelligibility of the standardized clinical word lists under varying degrees of time compression, provide a basis for comparison with future studies involving hearing-impaired listeners. The articulation scores generated by the normal listeners examined in this study were found to increase at the rate of 2-3.5%/dB for all conditions of time compression. In addition, for all except the 70% time compression condition, the mean articulation curves were found to be curvilinear, reaching an asymptote at the sensation level of 32 dB.

No differences between right and left ears were found. Implications of this finding may have bearing upon current theories of cerebral dominance for speech with respect to monotic versus dichotic listening tasks.

Figure 5.--Average articulation scores for each of four test lists plotted by sensation level

Finally, the interlist comparisons of the Northwestern Auditory Test No. 6 revealed essentially no list differences at the 0% time compression condition, but for the other conditions of time compression, List I was generally found to be most difficult whereas List IV was found to be easiest.

LIST OF REFERENCES

Aaronson, D., Immediate recall of compressed speech. Unpub. manuscript, Harvard Univ., Center for Cognitive Studies: Boston (1966).

Aaronson, D., Temporal factors in perception and short term memory. Psychological Bulletin, 67, 130-144 (1967).

Beasley, D. S., Auditory analysis of time-varied sentential approximations. Unpublished doctoral dissertation, University of Illinois (1970).

Bocca, E., Binaural hearing: another approach. Laryngoscope, 65, 1164-1171 (1955).

Bocca, E., Distorted speech tests. Sensorineural Hearing Processes and Disorders, A. B. Graham, ed., Boston: Little, Brown, and Co. (1967).

Bocca, E., Calearo, C., Cassinari, V., and Migliavacca, F., Testing cortical hearing in temporal lobe tumors. Acta Oto-Laryngologica, 45, 289-303 (1955).

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).

Borkowski, J. G., Spreen, O., and Stutz, J. Z., Ear preference and abstractness in dichotic listening. Psychonomical Science, 3, 547-548 (1965).

Broadbent, D. E., Immediate memory and simultaneous stimuli. Quarterly Journal of Experimental Psychology, 9, 1-11 (1957).

Broadbent, D. E., A mechanical model for human attention and immediate memory. Psychological Review, 64, 205-215 (1957).

Broadbent, D. E., and Gregory, M., Accuracy of recognition for speech presented to the right and left ears. Quarterly Journal of Experimental Psychology, 16, 359-360 (1964).

Bryden, M. P., Ear preference in auditory perception. Journal of Experimental Psychology, 65, 103-105 (1963).

Calearo, C., Binaural summation in lesions of the temporal lobe. Acta Oto-Laryngologica, 47, 392-397 (1957).

Calearo, C., and diMitri, T., Sulla intelligibilita della voce periodicemente alternata da un orchio all altro. Congress di Audiologica. Padova, Italy (1958).

Calearo, C., Teatini, G. P.,and Pestalozza, G., Speech intelligibility in the presence of interrupted noise. Journal of Auditory Research, 2, 179-187 (1962).

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

Chaney, R. B., and Webster, J. C., Information in certain multidimensional sounds. Journal of the Acoustical Society of America, 40, 447-455 (1966).

Corsi, P., The effects of contralateral noise upon perception and immediate recall of monaurally-presented verbal materials. Unpublished Master's Thesis (1967).

Curry, F. K., A comparison of left-handed and right-handed subjects on verbal and non-verbal dichotic listening tasks. Cortex, 3, 343-352 (1967).

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).

Darwin, C. J., Unpublished doctoral dissertation, University of Cambridge (1969).

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

Dirks, D., Perception of dichotic and monotic verbal material and cerebral dominance for speech. Acta Oto-Laryngologica, 58, 73-80 (1964).

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

Fairbanks, G., Everitt, W., and Jaeger, R., Methods for time or frequency compression-expansion of speech. Translations I.R.E.-P.G.A., AU-2, 7-12 (1954).

Fairbanks, G., Guttman, N., and Miron, M., Effects of time compression upon the comprehension of connected speech. Journal of Speech and Hearing Disorders, 22, 10-19 (1957).

Fairbanks, G., and Kodman, F., Word intelligibility as a function of time compression. Journal of the Accoustical Society of America, 29, 636-641 (1957).

Finzi, A., Il comportamento della soglia di intellezione del giovani con poacusia percetteiva e dei presbiacusici verso tests audiometrici vocali sensibilizzati. Arch. It. di Otol., 67 (1956).

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

Glorig, A., A report of two normal hearing studies. Annals of Otology, Rhinology and Laryngology, 67, 93-112 (1958).

Hirsh, I. J., The Measurement of Hearing. New York: McGraw-Hill Book Co. (1952) p. 147.

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

Jerger, J., Auditory tests for disorders of the central auditory mechanism. Neurological Aspects of Auditory and Vestibular Disorders. W. S. Fields and B. R. Alford, eds. Springfield: C. C. Thomas (1964).

Jerger, J., Audiological manifestations of lesions in the auditory nervous system. Laryngoscope, 70, 417-425 (1960).

Katz, J., The use of staggered spondaic words for assessing the integrity of the central auditory nervous system. Journal of Auditory Research, 2, 327-337 (1962).

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).

Kimura, D., Some effects of temporal-lobe damage on auditory perception. Canadian Journal of Psychology, 15, #3, 156-165 (1961).

Kimura, D., Functional assymetry of the brain in dichotic listening. Cortex, 163-178 (1967).

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

Mackworth, J. F., Paced memorizing in a continuous task. Journal of Experimental Psychology, 58, 206-211 (1959).

Miller, G. A., Language and Communication. New York: McGraw-Hill Book Co. (1951) p. 77.

Miller, G., and Licklider, J., The intelligibility of interrupted speech. Journal of the Acoustical Society of America, 22, 167-173 (1950).

Miron, M., and Brown, E., Stimulus parameters in speech compression. Journal of Communications, 18, 219-235 (1968).

Murphy, E. H., and Venables, P. H., Ear asymmetry in the threshold of fusion of two clicks: A signal detection analysis. Quarterly Journal of Experimental Psychology, 288-300 (1970).

Pollack, I., Message uncertainty and message reception. Journal of the Acoustical Society of America, 31, 1500-1508 (1959).

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

Rintelmann, W. F. and Schumaier, D., Personal communication (1971).

Shankweiler, D. P., and Studdert-Kennedy, M., Identification of consonants and vowels presented to left and right ears. Quarterly Journal of Experimental Psychology, 19, 59-63 (1967).

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).

Speaks, C., and Jerger, J., Method for measurement of speech identification. Journal of Speech and Hearing Research, 8, 185-194 (1965).

Sperling, G., The information available in brief visual presentations. Psychological Monographs, 74 (11, Whole No. 498) (1960).

Sperling, G., A model for visual memory tasks. Human Factors. 5, 19-31 (1963).

Sternberg, S., Two operations in character recognition: Some evidence from reaction time measures. In, Models for the perception of speech and visual form. AFRCL Symposium, Boston (1964).

Thorndike, E. L., and Lorge, I., Teacher's Word Book of 30,000 Words. New York: Bureau of Publications, Teachers 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, R., 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.

Willeford, 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).

APPENDIX A

FOUR LISTS OF FORM B

OF NU AUDITORY TEST NO. 6

N. U. AUDITORY TEST #6
N. U. RESEARCH P.B. LIST I FORM B

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. tip
40. page
41. raid
42. raise
43. bean
44. hash
45. limb
46. third
47. jail
48. knock
49. whip
50. met

N.U. AUDITORY TEST #6
N.U. RESEARCH P.B. LIST II FORM B

1. live
2. voice
3. ton
4. learn
5. match
6. chair
7. deep
8. pike
9. room
10. read (to read)
11. calm
12. book
13. dab
14. loaf
15. goal
16. shack
17. far
18. witch
19. rot
20. pick
21. fail
22. said
23. wag
24. haze
25. white
26. hush
27. dead
28. pad
29. mill
30. merge
31. juice
32. keg
33. gin
34. nice
35. numb
36. chief
37. gaze
38. young
39. keep
40. tool
41. soap
42. hate
43. turn
44. rain
45. shawl
46. bought
47. thought
48. bite
49. lore
50. south

N.U. AUDITORY TEST #6
N.U. RESEARCH P.B. LIST III FORM B

1. sheep
2. cause
3. rat
4. bar
5. mouse
6. talk
7. hire
8. search
9. luck
10. cab
11. rush
12. five
13. team
14. pearl
15. soup
16. half
17. chat
18. road
19. pole
20. phone
21. life
22. pain
23. base
24. mop
25. mess
26. germ
27. thin
28. name
29. ditch
30. tell
31. cool
32. seize
33. dodge
34. youth
35. hit
36. late
37. jug
38. wire
39. walk
40. date
41. when
42. ring
43. check
44. note
45. gun
46. beg
47. void
48. shall
49. lid
50. good

N.U. AUDITORY TEST #6
N.U. RESEARCH P.B. LIST IV FORM B

1. rose
2. dog
3. time
4. such
5. have
6. mob
7. bone
8. sail
9. rough
10. dip
11. join
12. check
13. wheat
14. thumb
15. near
16. lease
17. yearn
18. kick
19. get
20. lose
21. kill
22. fit
23. judge
24. should
25. pass
26. back
27. hall
28. bath
29. tire
30. peg
31. perch
32. chain
33. make
34. long
35. wash
36. food
37. mood
38. neat
39. tape
40. ripe
41. hole
42. gas
43. came
44. vote
45. lean
46. red
47. doll
48. shirt
49. sour
50. wife

APPENDIX B

INSTRUCTIONS GIVEN TO LISTENERS

INSTRUCTIONS GIVEN TO LISTENERS

You will now hear a tape recording of four lists, each composed of fifty monosyllabic words. Each word is preceded by a carrier phrase, "You will say". Your task will be to write down the word immediately following the carrier phrase in the appropriate space provided on the answer sheet. For example, if you hear, "You will say dog", you would be expected to write the word "dog". There will be an ample amount of time provided immediately after each word presentation for you to write down your response.

Each of the four lists will be presented to you at a different intensity level, although all of the fifty words on the same list will be equally loud. Some of the lists may sound extremely soft, so it is of extreme importance that you pay careful attention to the listening task. In addition, it may seem that the words are spoken on this tape in an unusually rapid manner, so again, pay close attention to what you hear and respond to the best of your ability. If you are uncertain of a response item, you are encouraged to guess. When you have completed an entire word list, there will be approximately twelve seconds before the items from the next list will be presented. Are there any questions?

APPENDIX C

ANSWER FORM USED BY THE LISTENERS

PR-5923-787-P.PAM
75-54T

No. ____________ Form ____________

Name __________________________ Sex ____________

Age ________ yrs. ________ mos. Date ____________

1 __________
2 __________
3 __________
4 __________
5 __________
6 __________
7 __________
8 __________
9 __________
10 __________
11 __________
12 __________
13 __________
14 __________
15 __________
16 __________
17 __________
18 __________
19 __________
20 __________
21 __________
22 __________
23 __________
24 __________
25 __________
26 __________
27 __________
28 __________
29 __________
30 __________
31 __________
32 __________
33 __________
34 __________
35 __________
36 __________
37 __________
38 __________
39 __________
40 __________
41 __________
42 __________
43 __________
44 __________
45 __________
46 __________
47 __________
48 __________
49 __________
50 __________