KRUG, Richard Fred. EFFECTS AND INTERACTIONS OF VISUAL AND AUDITORY CUES IN ORAL COMMUNICATION.

The University of Oklahoma, Ph.D., 1960
Speech – Theater

University Microfilms, Inc., Ann Arbor, Michigan

THE UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

EFFECTS AND INTERACTIONS OF VISUAL AND AUDITORY CUES IN ORAL COMMUNICATION

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

degree of

DOCTOR OF PHILOSOPHY

BY

RICHARD FRED KRUG

Norman, Oklahoma

1960

EFFECTS AND INTERACTIONS OF VISUAL AND AUDITORY

CUES IN ORAL COMMUNICATION

APPROVED BY

John W. Keys

P. T. Teska

Donald T. Cornish

Tom W. Tillman

Roger E. Nebergall

Kenneth H. Shank

DISSERTATION COMMITTEE

ACKNOWLEDGEMENTS

Acknowledgement is made to Dr. Tom W. Tillman for his assistance in the completion of the study. Gratitude is expressed to Dr. John W. Keys, director of this study, for his criticism and understanding throughout the entire course of this research.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

EFFECTS AND INTERACTIONS OF AUDITORY AND VISUAL CUES IN ORAL COMMUNICATION

CHAPTER I

INTRODUCTION

Although it is true that several senses are used to keep man in communication with his environment, the sensory avenues utilized in receiving oral communication are audition and vision. Auditory impressions or patterns refer to what is heard, and include not only those meanings consciously perceived, but also many subliminal cues which result from having heard the speaker. Subliminal cues, such as the pitch and inflection of the voice, the pauses between words, and the rhythm of speech, have something to add, over and above that which is being communicated by the words alone.

The visual cues utilized in oral communication may include an observation of the speaker's oral movements, body attitude, gestures, facial expression, and of general situational cues. In addition, subliminal cues which result from having visually observed the speaker may add meaning to verbal communication. The customary meaning of a word can be altered by a simple gesture, which may give a different nuance to the word meaning. The non-auditory messages received by the listener may reinforce, modify, or even contradict the words the individual has spoken.

Individuals working with the acoustically handicapped must, therefore, concern themselves with these two components of oral communication. With the hard of hearing, as well as with normal hearers, the ability to understand the speaker is related to the degree to which the available visual and auditory cues of speech are utilized.

Conventional lipreading tests evaluate the individual's ability to make use of only the visual cues of speech; auditory stimulation is absent. Conversely, tests of auditory discrimination evaluate the ability to make use of only the auditory cues of speech; in these instances the visual stimuli are unavailable to the listener.

For the profoundly deaf, therefore, speechreading tests are a relatively reliable index of the ability to understand speech. However, the same is not true for the hard of hearing. Compared with the deaf, the hard of hearing receive a greater amount of auditory stimulation. Consequently, the reception of speech for the hard of hearing is an audio-visual process, and it becomes apparent that neither lipreading tests nor tests of auditory discrimination, used singly, provide an accurate assessment of the ability of the hard of hearing to understand speech.

An evaluation can be made, then, of a person's ability to make use of the pure auditory component of speech (auditory discrimination scores). Estimates can also be made of the ability to make use of the purely visual aspects of speech (speechreading scores). It is of considerable importance to learn how the intelligibility of speech is increased as a result of combining the visual cues with the auditory cues. The effect of combining the auditory and visual components of speech is not easily seen, since the intensity of the auditory stimulus reaching the listener varies from moment to moment. At low intensity levels speech may be entirely unintelligible, while at higher levels it may be completely intelligible, even without visual cues. In studying bisensory reception of speech, it would be important to note how the role of visual stimulation changes as a function of the intensity level of speech.

Furthermore, the improvement in intelligibility which results from combining visual and auditory stimuli may vary with the individual. Some individuals may depend more upon the visual cues of speech than others. It appears reasonable to assume that those who must depend upon visual cues for understanding speech may have learned to utilize them more effectively. This would suggest that hard of hearing individuals, who are more dependent upon vision in communication, may make more use of the visual component of speech than do normal hearing individuals. It would be of interest, therefore, to compare the normal hearing with the hard of hearing in regard to the degree of improvement in the intelligibility of speech resulting from the addition of visual cues.

When words presented at a given intensity level are difficult to distinguish on the basis of auditory patterns alone, a premium may be placed upon the help afforded by the visual components. Conversely, when words are easily identified by auditory patterns alone, visual stimuli would not appear to be as important. It would seem logical, in this light, to ask if the contribution of visual cues varies according to the type of speech material presented.

Speechreading ability is usually taken to mean the ability to identify speech samples from visual cues in the absence of auditory stimulation. If the ability to make use of the visual cues of speech in the absence of the auditory patterns of speech differs from that ability when the auditory cues are available to the listener, then we may need to re-examine some of the techniques used in teaching speechreading to the hearing impaired. At the present time, many instructors insist that speechreading drill be carried on without voice, while others insist that all speechreading exercises be given with voice. The use made of the visual cues of speech in the absence of auditory stimulation may be one phenomenon, while the use made of the visual cues in the presence of auditory stimuli may be something altogether different. If this is true, instructors of speechreading may need to re-evaluate their stand regarding the use of voice in lipreading instruction or drill.

It is also true that if the ability to make use of the visual cues of speech varies with the presence or absence of the auditory component, there may be ramifications related to the recommendation of hearing aids for the severely hard of hearing and deaf. Hearing aids may provide the severely hard of hearing and the deaf with distorted, fragmentary, incomplete, or minimal auditory patterns of speech. These bits of auditory information may then supplement the visual cues available to the deaf, resulting in a considerable degree of improvement in the understanding of speech. Thus, it is altogether possible that the hearing aid should be considered an adjunct to lipreading, and its effectiveness evaluated from the point of view of providing the listener with the tools with which to obtain a maximum understanding of speech. It appears important, therefore, not only to evaluate independently the role of visual cues and the role of audition in the understanding of speech, but also to study their combined effect upon the intelligibility of speech.

In an effort to clarify the relative importance of visual cues and auditory cues in the intelligibility of speech and to assess the possible interaction of these two components, the present study was designed. It was designed to explore the following specific questions:

1. How much does speech intelligibility improve as a result of combining the auditory and visual stimuli as compared to performance when the stimuli are only available independently?
2. How does the contribution of visual cues to the intelligibility of speech vary as a function of the type of speech material presented to the listener?
3. What effect does hearing acuity and configuration of hearing loss have upon the use made of bisensory cues of speech?
4. How does the contribution of visual cues vary as a function of the intensity level of the auditory cues with which they are combined?

To answer these questions, three groups of subjects were used. One group had normal hearing, a second group had moderate hearing losses of a relatively flat configuration, and a third group consisted of persons with moderate hearing losses with a sloping configuration. Each subject responded to two types of word lists: phonetically balanced word lists and a multiple choice intelligibility test. A word list of each type was presented to each subject at eight auditory sensation levels ranging from -6 to +15 db. At each sensation level, the experimental subjects responded to speech samples under two conditions: an auditory and a bisensory presentation of a word list. Under the auditory presentation, only the auditory patterns of speech were available to the listener, while under the bisensory presentation, both the auditory and visual cues of speech were available. A third condition, a purely visual presentation, provided the listener with only the visual component. It would be expected that an evaluation of the differences in visual, auditory and bisensory performances among subjects with varying degrees of hearing acuity and with different types of audiometric configurations would yield a greater understanding of the relative importance of the visual and auditory components of speech and of the interactions between them.

CHAPTER II

HISTORY

When working with acoustically handicapped individuals, the teacher or clinician often places a great deal of emphasis upon the ability of an individual to make use of the visual cues of oral communication. If the person possesses a small degree of residual hearing, communication may be bisensory even though it may be predominantly visual. As important as visual cues appear to be, there is only a limited amount of research which compares the auditory reception of speech with the bisensory (audio-visual) reception of speech of acoustically handicapped individuals. The research reported in the literature will be reviewed at this time.

One of the earliest studies comparing the auditory reception of speech to an audio-visual reception is reported by Ewing.[1] In that study, intelligibility scores, based upon the correct recognition of sentences, were obtained under various conditions for 92 adventitiously hard of hear-

---

[1] I. Ewing, Lipreading and Hearing Aids (Manchester: Manchester University Press, 1946).

ing subjects ranging in age from 17 to 72 years. Neither the range of severity of hearing loss nor the average hearing loss for the group was reported. Under ordinary listening conditions, without lipreading and without the use of a hearing aid, the average intelligibility score for the group was 21% correct. The type of speech material used and the intensity level at which the speech material was presented were not reported. When speechreading was used simultaneously with residual hearing, the group correctly identified 64% of the speech samples. When speech was amplified with a three-stage hearing aid, and the subjects prohibited from visually observing the speaker, a score of 64% was obtained. Under the final condition of listening with a three-stage hearing aid, together with speechreading, a score of 90% was earned. These results indicate that both speechreading and the amplification of sound, when accomplished independently of each other, enhance the intelligibility of speech. In addition, there seems to be a fortunate interaction between these two components of verbal communication which allows for maximum understanding of speech when they are presented simultaneously.

From these findings, it appears that the acoustically handicapped make use of both the visual and auditory cues of speech. Although it is true that normal hearing individuals make much use of the auditory component of speech, the degree to which they are able to utilize the visual cues of speech may be a matter of speculation. O'Neill, who was aware of the importance of both the visual and auditory components of speech states:

> Speech is usually regarded as an oral-auditory process. It has visual characteristics, however, that the deaf and hard of hearing may employ in the understanding of speech through lipreading, and that may contribute to normal communication. Since most verbal communication is direct, face-to-face, oral sending-receiving, the perception of speech might be regarded as a bisensory (auditory-visual) phenomenon.[1]

As a result of his interest in the problem, O'Neill assessed the relative contribution that speechreading makes in person-to-person communication using vowels, consonants, words, and phrases as speech materials. Thirty-two subjects listened to an auditory presentation and an audio-visual presentation of speech under four speech-to-noise ratios. Under the audio-visual condition, the subjects were allowed to view the speaker as he presented word lists at the four speech-to-noise ratios. Under the auditory condition, the subjects were not permitted to observe the speaker as the word lists were presented. The speech-to-noise ratio was described as the difference in decibels between the intensity of the speech signal and the intensity of the noise. A speech-to-noise ratio of 0 db indicated that both stimuli were of the same intensity, while a positive speech-to-noise ratio indicated that the intensity level of the speech signal

---

[1] John J. O'Neill. "Contributions of the Visual Components of Oral Symbols to Speech Comprehension," Journal of Speech and Hearing Disorders, XIX (1954), 429.

was greater than that of the noise. Conversely, a negative speech-to-noise ratio indicated that the intensity of the speech was less than that of the noise.

The noise levels employed were 36 db (-20 db speech-to-noise ratio), 76 db (-10 db speech-to-noise ratio), 66 db (0 db speech-to-noise ratio), and 56 db (10 db speech-to-noise ratio). All speech and noise levels were measured re .0002 microbar. The noise was uniform spectrum random noise with frequencies above 10,000 cps filtered out. The mean speech level, as controlled by visual monitoring with the aid of a magnetic throat microphone leading to a voltmeter, was 66 db when measured at a distance of two feet in front of the speaker. The subjects were normal hearing undergraduate students grouped into four listening panels of eight members each. Each panel was seated in a semi-circle eight feet from the speaker.

There was no report of the use of earphones, nor was it specified that testing was performed under free-field conditions. However, since speech level measurements were made at a distance of two feet from the speaker, and since a panel of eight listeners was tested at one time, it is assumed that all tests were made under free-field conditions. A basic assumption of the study was that the score obtained at the -20 db speech to noise ratio (when the listener viewed the speaker), would indicate the speechreading score of a visually observed oral stimulus. Obtained in this manner, the mean speechreading scores were 44.5% for vowels, 72% for consonants, 64.1% for words and 25.9% for phrases. However, the test results also showed that the contribution of visual cues to the total number of words which were correctly identified decreased as the speech-to-noise ratio varied from -20 db to 10 db. For example, at the -20 db speech-to-noise ratio (speech below noise level), the mean auditory word recognition score (number of words correctly recognized) was 5.44%, and the mean audio-visual (bisensory) scores 13.4%. This improvement at the -20 db speech-to-noise ratio can be compared to that at the 10 db speech-to-noise ratio (speech level higher than noise level), where the mean auditory score was 17.5% and the bisensory score 18.6%. O'Neill then mentions that:

> It appears, in effect, that when vision and audition jointly contribute to the amount of information received, they supplement each other. But the combined effectiveness of the two, when quantified, is less than the summation of their individual contributions to recognition. From the vocabulary of information theory, another description, or explanation, of this is that there is a redundancy in the recognition of sensory signals; that each sensory modality is contributing similar information.[1]

It is not clear why the investigator, when summarizing his findings, considered the mean recognition score at the -20 db speech-to-noise ratio as the contribution of visual cues to speech in view of his report of differences in the contribution of visual cues as a function of the

---

[1] Ibid. 434.

speech-to-noise ratio.

The virtue of considering conventional laboratory articulation tests as indications of the person's performance in many practical work situations was questioned by Sumby and Pollach.[1] They felt that the standard laboratory articulation tests may be misleading as a result of several factors, two of which were: "...the information associated with the class of possible messages and the contribution of visual factors to speech intelligibility."[2] First, if only a small number of possible messages may be communicated, higher noise interference levels can be tolerated than if the class of possible messages is large. Secondly, if visual cues of speech supplement the auditory patterns, a higher noise interference level can be tolerated than if visual cues are not utilized. The experiment they conducted considered the interaction of these two factors. More specifically, they examined the contribution of visual cues to speech intelligibility as a function of the speech-to-noise ratio and the size of the possible vocabulary.

In that study, the speech samples employed were 256 spondaic words randomly classified into six groups. Each group contained a different number of spondaic words and

---

[1] W. H. Sumby and I. J. Pollach, "Visual Contribution to Speech Intelligibility," Journal of the Acoustical Society of America, XXVI (1954) 212-215.

[2] Ibid. 212.

became known as a "vocabulary class." Thus, vocabulary classes of 8, 16, 32, 64, 128 and 256 words were established. From each of these vocabulary classes, test lists of from 25 to 50 words were constructed. In addition, tests of monosyllabic words and trisyllabic phrases were developed, with the trisyllabic phrase consisting of a spondaic word plus a monosyllabic word. The speech samples were binaurally presented to normal hearing subjects by monitored live-voice through earphones. The noise, presented simultaneously with the speech, was uniform in level per cycle in the frequency range of 20-10,000 cps. To attain the speech-to-noise ratios, which ranged from -30 db to 0 db, the noise level was held constant and the speech level varied. The noise level was not reported. Prior to presenting each list, the speaker recited the test vocabulary in order to define the words under test, and a printed reference list of the vocabulary under test was then given each subject. The subject selected his response from the printed restricted vocabulary. Performance was evaluated under two test conditions: first, when only auditory cues were available, and secondly, when both the auditory and visual cues of speech were presented to the listener. The results of the study can be summarized as follows:

1. Under the condition of listening only (no visual cues), the intelligibility of speech increased as the speech-to-noise ratio decreased from -30 db (speech 30 db lower than noise level) to 0 db

(speech and noise at the same intensity level). For example, with the 8-word vocabulary class, the approximate increase was from 15% to 98% correct.

2. Under conditions of listening only, the intelligibility of speech at a given signal-to-noise ratio decreased as the size of the vocabulary class increased from 8 words to 32 words. Little change in the intelligibility of speech was observed as the size of the vocabulary class was increased beyond 64 words. For example, at the -18 db speech-to-noise ratio, intelligibility scores decreased from 38% for the 8-word vocabulary class, to 15% for the 64, 128 and 256-word vocabulary classes.

3. Under the bisensory condition (auditory and visual cues available), the intelligibility of speech also varied as a function of the speech-to-noise ratio and the size of the vocabulary in much the same manner as it had under the auditory presentations. The striking difference, however, was in the higher resistance to noise interference for the bisensory presentations. As an example, at a speech-to-noise ratio of -30 db, the score for the 8-word vocabulary class was 15% correct when only auditory cues were available, as compared to an intelligibility score of 90% for the bisensory presentations. Furthermore, at the speech-to-noise ratio of -24, the per-

centage correct scores decreased from 98% for the 8-word vocabulary class to 44% for the 128-word vocabulary class.

4. Under a noise-free condition (the level of speech not reported, but assumed to be 66 db re .0002 microbar), there was no appreciable difference in intelligibility scores for the two test conditions (auditory vs. bisensory). Furthermore, in the absence of noise, the size of the restricted vocabulary did not materially affect the intelligibility scores.
5. For the bisensory presentations, no real variation in intelligibility scores resulted as a function of the word length.
6. For the auditory presentations, higher scores were obtained for the bisyllabic and trisyllabic vocabularies than for the monosyllabic vocabulary.

The study was a descriptive one, and no attempt was made to evaluate the results in statistical terms. However, it clearly pointed out that normal hearing subjects can utilize the visual component of speech to increase the understanding of verbal communication.

Noting the work of Sumby and Pollach,[1] Neely[2]

---

[1]Sumby and Pollach, op. cit.

[2]K. K. Neely, "Effects of Visual Factors on the Intelligibility of Speech," Journal of Acoustical Society of America, XXVIII (1956), 1275-1277.

emphasized the fact that many times voice communication must be carried on in high intensity noise environments. Moreover, since the intelligibility of speech is a direct function of the speech-to-noise ratio, high intensity noise tends to lower the intelligibility of speech heard in such a background. It was his feeling that the knowledge of how the intelligibility of speech is affected by combining the visual cues with the auditory stimuli would be extremely valuable in developing or assessing voice communication procedures. Sumby and Pollach[1] evaluated the effect of visual cues in high noise level environments, while Neely[2] was interested in evaluating the effect of visual cues on listener intelligibility in terms of distance and angle from which the listener observed the speaker.

The 35 normal hearing males used by Neely[3] in the study listened to the multiple choice intelligibility lists developed by Black.[4] The response to each of the 27 test items was selected by the listener from a choice of four words. Speech samples were presented to the listeners with monitored live-voice through earphones, while the noise, generated by a random-noise generator, was introduced into

---

[1]Sumby and Pollach, op. cit.

[2]Neely, op. cit.

[3]Neely, op. cit.

[4]J. W. Black, "Multiple-Choice Intelligibility Tests," Journal of Speech and Hearing Disorders, XXII (1957), 213-235.

the test room through loudspeakers at an overall level of approximately 100 db re .0002 microbar. The noise level was high enough to ensure that the only auditory speech patterns heard by the listeners were those received through the earphones. At the ear, the pressure level of speech was approximately 80 db re .0002 microbar. The speaker, together with the listeners, were in a test room where all listeners could be tested at the same time. To allow each person to receive each treatment, subjects rotated positions. Under the first three test conditions, the subjects sat at an angle of 90 degrees to the speaker, and at distances of 3, 6, and 9 feet from the speaker. For the second three test conditions, the listeners sat at an angle of 45 degrees, and at distances of 3, 6, and 9 feet, while in the three final test positions, the listeners directly faced the speaker at the distances of 3, 6, and 9 feet. The results, subjected to a statistical analysis, indicate that:

1. The bisensory presentations resulted in significant increases in word intelligibility when compared with performance under a pure auditory presentation of speech.
2. The angle from which the listener observed the speaker had an influence on intelligibility scores. For instance, at a distance of 9 feet, the mean intelligibility score for the group was 59% when the speaker was observed at an angle of 90 degrees,

while at the 45 degree angle, the mean score was 62%. A further, although slight increase to 64% was recorded under the face-to-face condition.

3. The distance from listener to speaker did not have a significant effect upon the intelligibility scores within the range of 3 to 9 feet.

It was concluded from the results, that the amount which visual cues contribute to the intelligibility of speech is dependent, to some extent, upon the angle from which the listener observes the speaker.

The research evaluating the contribution of visual cues to speech discrimination was reviewed by Hutton,[1] who pointed out that there was little information on how the discrimination of individual speech sounds was affected by combining the visual cues with the auditory patterns. He then designed a study enabling him to obtain intelligibility scores on consonants and vowels under three conditions. Under the first, a visual condition, only the visual cues of speech were available to the listener. The speaker was observed by the subject through a double glass window which connected a suite of two sound treated testing rooms. Under the second, or the auditory condition, only the auditory patterns of speech were available to the listener. The auditory stimuli were presented by monitored live-voice by means

---

[1]Charles Hutton, "Combining Auditory and Visual Stimuli in Aural Rehabilitation," Volta Review, LXI (1959), 316-319.

of a speech audiometer and were received by the listener through his own hearing aid or an auditory training system. It was not reported whether the auditory stimuli received through the auditory training unit were received binaurally or monaurally by the listener, although it is assumed that for subjects using their own hearing aids, speech was received monaurally. No mention was made of attempts to control or measure the intensity of speech reaching the ear of the subjects. Under the third, or bisensory condition, both the visual and auditory cues of speech were available to the listeners. The subjects listened through their own hearing aids or through the auditory training unit and simultaneously observed the speaker through the double glass window connecting the suite of two testing rooms.

Speech materials were multiple-choice word lists, each item consisting of four words exhibiting as little phonemic contrast as possible. Twenty-five rehabilitation patients with mild to profound hearing losses served as subjects. In presenting the test items to the listeners, several different speakers were used. The following is a summary of the findings.

1. When the consonants were evaluated, the mean value (percentage of words correct) indicated that, as a whole, about 50% more information was received from auditory stimuli than from visual stimuli. The mean score of the bisensory presentations was approxi-

mately double the mean of the visual scores. However, the bisensory mean score was considerably less than the sum of the auditory and visual mean scores.

2. There were large differences among the various phonemes in intelligibility gains as a result of combining stimuli. For instance, the intelligibility of [b] and [v] benefited substantially from the addition of visual cues to the auditory patterns, whereas, with the [l] and [m], little benefit was seen from combining the stimuli.
3. There was no apparent relationship between the visual intelligibility scores and the degree of improvement in intelligibility resulting from the bisensory reception of speech.
4. It was found that the mean auditory intelligibility scores for vowels was considerably higher than that of the consonants. However, the visual score means for the vowels and consonants were about the same. This was felt to indicate that auditory stimuli played a much larger role in the bisensory recognition of vowels than they did in the bisensory recognition of consonants.

The study suggests that the intelligibility of phonemes varies depending upon the sensory avenue utilized in their reception.

It must be pointed out again, however, that the level of speech presented to the subjects was not mentioned. Moreover, the amount of gain afforded by the auditory training unit or the hearing aid was not considered. Finally, several different speakers were used in testing the subjects. Each of the three mentioned variables would have an effect upon the intelligibility of speech, but did not seem to be controlled adequately in the study.

In brief, there are relatively few studies which attempt to compare auditory, visual, and bisensory word recognition scores; thus, the significant findings provided by the studies mentioned can be summarized in the following manner:

1. The addition of the visual cues of speech to the already present auditory stimuli enhances word intelligibility scores. The improvement occurs whether the listener is in a noise-environment or in a relatively quiet environment.
2. In a noise-environment, the relative intensity of speech has an influence upon the contribution of visual cues to word intelligibility. In general, as the level of speech is increased in relation to the noise level, the contribution of visual cues decreases.
3. In a noise-environment (-20 db speech-to-noise ratio), speechreading contributes most to the understanding of consonants, and less to the understanding of words, vowels, and phrases, in the order named.

4. In a noise-environment, the angle from which the listener observes the speaker affects the word recognition scores. The greater the angle from which the face of the speaker is observed, the poorer are the visual word recognition scores.
5. When the listener is in a noise-environment, the size of the vocabulary from which he must choose a response affects the word recognition score. In general, as the size of the vocabulary increases, the intelligibility of speech decreases. However, when speech is presented at a moderately intense level, and the listener is in a relatively quiet environment, the size of the vocabulary from which he must choose his response has no effect upon the intelligibility of speech.
6. In a quiet-environment, vowel sounds are more easily identified than are the consonants through auditory cues alone.
7. In a quiet-environment, all phonemes do not benefit equally (in terms of increased intelligibility) from a bisensory mode of presentation.
8. There is no apparent relationship between speech-reading ability and the degree of improvement in intelligibility when the visual cues are combined with the auditory cues of speech.

It is felt that the previous studies were relatively limited or wanting in one of more respects. For those studies employing hard of hearing subjects, the population in which we are most interested, the variables affecting the word recognition scores were not well controlled. The variables which did not appear well controlled included the intensity level of the speech presented to the subject, the severity of hearing loss among the subjects, and the degree of amplification obtained through the hearing aids or auditory training units.

In those studies employing normal hearing subjects, the variables appeared better controlled. However, the studies considered only the performance of normal hearing subjects, and used only one type of speech material.

It appears there is a need for further study of the relationship between the visual and auditory cues of speech. More specifically, since the intensity of speech varies continuously, closer attention must be paid to the changing relationship of the visual and auditory components of speech as a function of the intensity of the auditory signal. Furthermore, additional information is needed regarding the performance of normal hearing individuals as compared to that of hard of hearing individuals. In addition, a comparison of the performance of two groups of individuals with the same degree of hearing loss but with losses of differing configuration would add to the body of knowledge at hand. Finally, knowing that speech materials differ in terms of discriminatory difficulty, it is necessary to explore the problem with more than one type of speech sample.

The present study, then, is an attempt to control some of the variables affecting the intelligibility of speech and to obtain further information regarding the contribution of visual and auditory cues to speech intelligibility.

CHAPTER III

SUBJECTS, APPARATUS AND PROCEDURES

Introduction

Three groups of subjects were used in the present study. Group I consisted of eight subjects with normal hearing. Group II was composed of eight subjects with moderate air conduction hearing losses; losses for this group were relatively flat. Group III was also composed of eight subjects with moderate air conduction hearing losses, but in this group the subjects had poorer hearing for the higher frequencies than did those subjects in Group II.

Each subject responded to two types of speech samples. One was a pencil and paper multiple-choice test of word recognition, and the other a word recognition test which required that the subject merely repeat aloud what he considered to be the stimulus word. Both types of speech samples were presented under three experimental conditions. The first was an auditory recognition condition, in which only hearing was used in obtaining a recognition score. An auditory recognition score was the total number of words correctly recognized by the subject when a word list was presented without visual cues through an earphone. The second condition was a visual recognition condition, under which only the visual cues of speech were available. The visual recognition score was the number of correct responses made to the word lists when only the visual cues of speech were available. The third condition was bisensory, in which both hearing and vision were used. A bisensory recognition score was the total number of words correctly identified when a word list was presented with both visual and auditory cues.

Under both the auditory and bisensory presentations, word lists were presented at eight sensation levels: -6 db, -3 db, 0 db, 3 db, 6 db, 9 db, 12 db and 15 db. A word list of each type was presented at each intensity level, and auditory and bisensory scores were recorded at each intensity level.

Subjects

A total of twenty-four subjects living within the metropolitan area of Oklahoma City, Oklahoma, took part in the study. All subjects were adults who had acquired speech and language in the normal manner through hearing, and each used the spoken language as the primary means of communication. None had visual defects which would preclude adequate observation of the experimenter at a distance of eight feet. With a standard Snellen Eye Chart, each subject earned a binocular vision score of at least 20/20 as measured when the subject was placed 20 feet from the chart. In addition, each subject was asked to read a portion of a word list when the material was 12 to 15 inches from his eyes.

None of the subjects presented a history of cerebral hemorrhage, stroke or aphasia. Also, none had received psychiatric or psychological therapy, nor did any exhibit inconsistent responses to audiometric tests or display behavior patterns which would suggest the possibility of a non-organic hearing loss. Each subject was selected in accordance with the criteria established for one of the three groups of eight subjects used in the study.

Group I

Subjects in Group I had normal hearing acuity bilaterally. The criterion for normal hearing was the ability to hear air conducted pure tones bilaterally at all octave intervals from 125 to 8000 cps at a hearing level of 20 db, reference USPHS norm. It was felt that if a person could respond to pure tones at the 20 db level, there should be no difficulty in hearing everyday speech. The average speech reception threshold hearing level for the group was -6 db or 23 db re: .0002 microbar. The group contained two males and six females ranging in age from 22 to 34 years, with a mean age of 31.5 years.

Group II

Group II consisted of subjects having an air conduction hearing loss of moderate degree and showing a relatively flat audiometric configuration. A flat audiometric configuration was defined as one in which hearing acuity varied no more than 10 db among any of the pure tones in the speech range (500, 1000 and 2000 cps). The criteria established for Group II required that each subject exhibit a minimum hearing loss of 30 db in either ear, and a maximum loss of 65 db in the ear under test. The degree of hearing loss was specified as the average loss for the frequencies 500, 1000 and 2000 cps. The mean hearing level at each test frequency for this group is shown in Figure 1. The average hearing loss for the three frequencies in the speech range (500, 1000 and 2000 cps) was 46 db. Among the subjects, the average loss for pure tones in the speech range varied from 32 db to 52 db. The mean speech reception threshold hearing level for the group was 45 db.

Duration of hearing loss among subjects ranged from 10 years to 50 years, with an average duration for the group of 23.9 years. Three subjects in the group possessed conductive-type hearing losses, three had perceptive-type losses, and two had mixed-type hearing losses. The criterion for a conductive-type loss was that bone conduction losses not exceed 10 db at more than one frequency in the speech range. The criterion for a predominately perceptive loss was that the bone conduction acuity be no more than 10 db better than the air conduction acuity for two of the three frequencies in the same speech range. Subjects not meeting

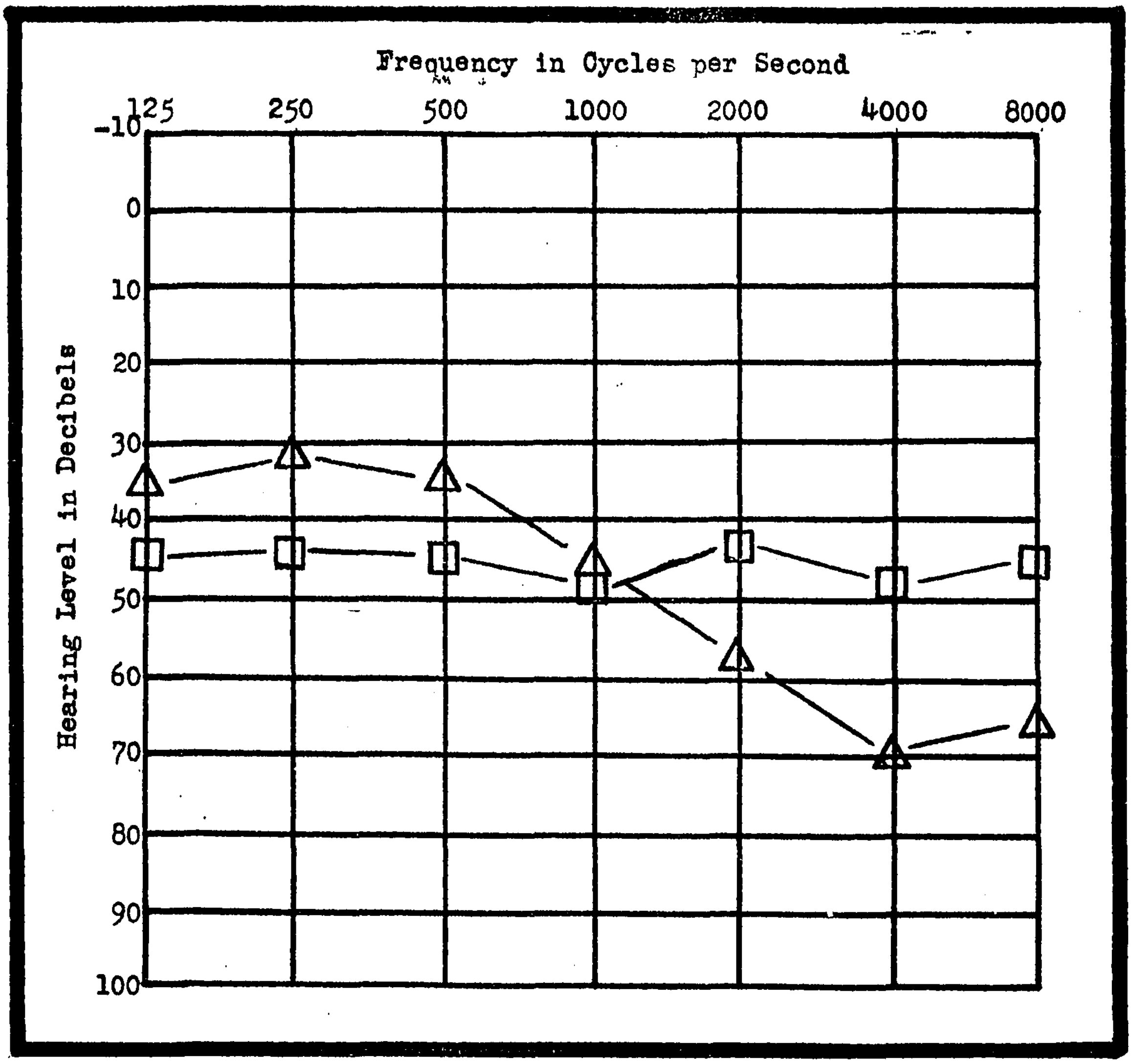

Figure 1. Mean hearing level in decibels re: USPHS norm for Group II (□) and Group III (△).

either of the above criteria were considered to have mixed perceptive-conductive hearing losses.

All eight subjects of Group II were females ranging in age from 36 to 69 years of age, with the group having a mean age of 46.6 years. Four wore hearing aids, and three had received a total of approximately eight hours of speechreading instruction during the two months immediately preceding the testing. Of the eight subjects, two felt they could speechread adequately, four felt they obtained only a limited amount of help, and two felt they received no help from speechreading.

Group III

Group III contained subjects also having air conduction hearing losses of moderate degree, but with sloping audiometric configurations. The criteria established for the group required that each subject have an air conduction hearing loss of at least 30 db in the better ear, the degree of hearing loss being the average of the pure tones in the speech range. It was further required that each subject have an air conduction hearing loss not greater than 65 db in the ear to be used in the word recognition test, as averaged from the pure tone losses at 500, 1000 and 2000 cps. Finally, the criteria demanded that the hearing loss at 2000 cps be at least 15 db greater than the hearing loss at 500 cps.

Thus, the subject selection criteria for Group III differed from that of Group II only in terms of the shape of the audiometric configuration. Subjects in Group II had relatively flat audiometric configurations, while those in Group III had a sloping configuration. The mean hearing loss for the group for all frequencies tested is shown in Figure 1. Among subjects of Group III, the average loss for pure tones in the speech range varied from 30 db to 65 db. For the group the average hearing loss for pure tones in the speech range was 46 db and the mean hearing loss for speech was 38 db.

Duration of hearing losses ranged from 2 years to 40 years, with an average duration of 15.8 years. As evaluated by the criteria established for type of hearing loss, seven subjects in the group had perceptive-type losses and one had a mixed-type, predominantly perceptive. Four of the subjects were males and four were females. The age range was from 26 years to 58 years, with a mean age of 46 years. Only two of the subjects wore hearing aids, and none had received speechreading instruction. Of the eight subjects, only one felt he could speechread adequately, five felt they obtained only limited help, and two felt they received no help from speechreading.

Apparatus

Test Rooms

All testing was conducted in the sound-proofed rooms of the Speech and Hearing Clinic of the University of Oklahoma. A suite of two rooms was employed, with one room serving as the test room for the subjects and the other as the control room for the experimenter. A three-paned soundproofed window just above the instrument panel of the speech audiometer allowed visual communication between the rooms. The window (15" x 38") did not distort vision or adversely affect the reception of the visual cues of speech. A 100-watt lamp was situated so as to reflect light off the wall in front of the experimenter, thus eliminating the slight shadow created by the overhead lighting.

Audiometric Equipment

The initial pure tone tests used to select subjects for groups was performed with a commercially available pure tone audiometer (Audio Development Company, Model 53C) feeding either of two earphones (Telephonics 39-10Z) or a hearing aid type bone conduction oscillator. The earphones were mounted in MX-41/AR cushions and were held in a standard headband.

Before testing, the acoustic output of the air conduction system of this instrument was calibrated with an audiometer calibrating unit (Allison, Model 3A), to the USPHS norm. The calibration of the bone conduction unit was clinically evaluated by comparing the bone conduction thresholds to the air conduction thresholds of seven subjects who, according to clinical records and previous examination, were held to exhibit perceptive losses. The bone conduction thresholds of these subjects were consistently superimposed upon the air conduction thresholds, and since no consistent variation occurred between the two thresholds, the bone conduction circuit was considered adequately calibrated for the purposes of this study.

All speech samples were presented by monitored live-voice, using a partially transistorized custom built four-channel speech audiometer especially designed for use in the audiology laboratories of the University of Oklahoma. The output of the speech audiometer was adjusted to produce a 29 db SPL re: .0002 microbar, with the hearing level of the audiometer set to zero. The acoustic output was transduced by either of two earphones (Telephonics, Type 39-10Z) housed in MX-41/AR cushions.

Procedures

Pre-testing Procedures

Before testing, it was necessary to select the speech materials to be used and to select the sensation levels at which the speech stimuli were to be presented. It was also necessary to randomize the presentations of the speech samples in such a way as to minimize the occurrence of consistent errors in the experiment.

Two types of word lists were selected for testing all subjects. The first was the Phonetically Balanced Word Lists (PB Lists) developed by the Psycho-Acoustic Laboratory of Harvard University.[1] The second type of speech sample was the Multiple-Choice Intelligibility Tests developed by Black.[2] Administration of the Black Lists requires the subject to choose a response from words which are auditorily similar and often confused with one another through hearing alone. Thus the effect of the addition of visual cues could be observed in instances where auditory discrimination between words was relatively difficult. The PB Lists presented the subject with the task of responding to a list of words with a wider range of auditory difficulty. Here the addition of visual cues could be observed in instances where the speech sample presented a range of auditory difficulty.

<u>Characteristics of the Phonetically Balanced Word Lists</u>.--Davis[3] considers the PB Lists one of the best tools for measuring auditory discrimination ability. There are 20 lists, each containing 50 monosyllabic words. Seventeen lists were used in the present study.[4] Each of the lists has the following characteristics:

1. All words are monosyllabic in structure and in common usage.

---

[1]J. P. Egan, "Articulation Testing Methods," *Laryngoscope*, LVIII (1948), 955-991.

[2]J. W. Black, "Multiple-Choice Intelligibility Tests," *Journal of Speech and Hearing Disorders*, XXII (1957), 213-235.

[3]H. Davis, (ed.) *Hearing and Deafness*. New York: Murray Hill Books, 1947.

[4]Egan, *op. cit*., lists 1 through 17.

2. Lists are of the same average difficulty, have an equal range of difficulty, and are of equal phonetic composition.

3. Each list includes all the common speech sounds of the English language in approximately the same proportion in which they occur in ordinary speech.

The PB Lists were selected as speech samples for several reasons. First, they are one of the few tests of discrimination having a sufficient number of lists of equal difficulty. Secondly, since the phonemes of the English language are represented in equal proportion from list to list, the visual cues of speech would appear also to occur in equal proportion from list to list. Finally, the administration of the PB Lists requires the subject to respond to words with a relatively wide range of auditory difficulty.

Characteristics of the Black Multiple-Choice Intelligibility Lists.--The multiple-choice lists, Forms C and D, developed by Black[1] constituted the second type of speech sample. In total, there are 24 multiple-choice lists, each list containing 27 multiple-choice test items. These lists were designed primarily to provide measurements for assessing military communication equipment. Each of the 17 lists[2] used in the study have the following characteristics:

---

[1]Black, op. cit.

[2]Black, op. cit., Lists used were: 1A, 5A, 6A, 8A, 10, 11A, 12A, 13, 13A, 14A, 15, 16, 17, 19A, 22A, 23 and 24.

1. Each word contains either one or two syllables and is in common usage.
2. Each word falls within a range of 15 to 85% intelligible in both quiet and in simulated propeller-type aircraft noise when presented by means of an aircraft intercommunication system.
3. The test items (words) are confused one with another on the basis of auditory cues alone.
4. On the scoring sheet, each test item contains four words from which the subject selects only one as his response.
5. Each of the lists is equated with respect to the mean level of difficulty of the test items.

The multiple-choice tests, Forms C and D, were selected because there were sufficient lists of equal average difficulty to allow each list to be used once for each of the 17 test conditions of this study. Furthermore, presentation of these lists required that the subject choose a response from words which were similar and often confused one with another through hearing alone. It was felt that these lists would provide an estimate of how visual cues aid in the correct recognition of words which are difficult to distinguish on the basis of auditory patterns alone.

Establishing the sensation levels for the presentation of word lists.--In order to establish the minimum and maximum sensation levels at which speech samples were to be presented to the subjects, two factors were considered. First, it was important that the minimum sensation level provide the subject with sufficient auditory cues to stimulate him to respond to the stimulus word. Secondly, it was necessary that the maximum sensation level for auditory presentations be low enough to prevent subjects from obtaining a perfect word recognition score. If, at maximum auditory sensation level subjects were allowed to earn perfect scores, there would be no opportunity for the visual cues of speech to improve word recognition scores. Consequently, each list type was presented at sensation levels of -6 db, -3 db, 0 db, 3 db, 6 db, 9 db, 12 db and 15 db.

Randomizing the presentations.--In order to minimize the effects of experimental errors, the experimental conditions and their order of presentation were randomized. The order of presentation according to type of word list was first randomized. Specific word lists were then assigned to the schedule of presentations for each type of word list. The condition of presentation (auditory cues or bisensory cues) was then assigned by chance. And finally, the sensation level at which each presentation was to be made was randomly assigned. The lists to be used for the two visual presentations were then randomly inserted into the order of testing. Only one visual presentation of each type of word list was made to each subject. Furthermore, the word lists used were the same for each subject. This was necessary to ensure that speechreading tests (visual presentations) were of equal visual difficulty.

The resulting testing order was called the "testing schedule." The testing schedule thus consisted of 34 randomized experimental treatments or presentations. Printed forms of the individual word lists were then arranged according to the testing sequence set forth by the testing schedule and were later used by the experimenter in presenting the word lists to the subject. In addition, the answer sheets for the miltiple choice tests were arranged in sequence for later use by the subject.

Preliminary Audiometric Procedures

Each subject received all tests and treatments individually, in a single experimental session. First, an air conduction test of hearing acuity was performed for each ear. Thresholds for pure tones were obtained with a standard commercial audiometer. Thresholds were established in the following manner: The 1000 cps tone was first presented at a hearing level of 40 db to 50 db. The subject was requested to raise his hand when he heard a tone. If the tone was not heard, the intensity was increased in 10 db steps until the subject indicated he did hear it. The tone was presented at the discretion of the experimenter, with short periods of silence between presentations of the tone. Following a response, the intensity was reduced in 5-db steps until the subject failed to respond. The intensity was then increased 10 db and decreased again in 5-db steps until the subject again failed to respond. The intensity was then increased 10 db and decreased again in 5-db steps until the subject again failed to respond. The intensity was once again increased 10 db, and for the third time decreased in 5-db steps until the subject once again failed to respond. The lowest intensity level eliciting two responses out of three presentations was taken as the threshold for the tone. The pure tones were tested in the following order: 1000 cps, 2000 cps, 4000 cps, 8000 cps, 500 cps, 250 cps and 125 cps.

Once the air conduction thresholds were established for each ear, bone conduction thresholds were obtained for the pure tones 250 cps, 500 cps, 1000 cps, 2000 cps, and 4000 cps. Bone conduction thresholds were obtained using the same technique employed in establishing air conduction thresholds. Masking through air conduction receivers was employed only when the test tone lateralized to the ear contra-lateral to the one under test, in which case the level of masking was raised only to the level required to return the test tone to the ear under test. In instances where the test tone could not be returned to the ear under test, the maximum masking level available on the audiometer was used.

Upon completing the pure tone tests, the subject was positioned in a tablet-arm chair in front of the sound-proofed window which separated the control room and test room. The experimenter then sat next to the subject and instructed him in the following manner:

> You are going to hear many lists of words. There will be two kinds of lists. With one kind of word list you need only repeat the last word I say. For instance, if I should say, 'Say the word house,' all you need to do is to repeat the word house. If I should say, 'Say the word cow,' all you need to do is to say the word cow. We will call these word lists the 'say the word lists.'
>
> The other kind of word lists we will call the 'scratch out lists,' because instead of saying the word you think I said, you merely scratch out the word you think I said. On the scratch out lists I will first say a number, and then three words, like this: 'Number One, virtue, blend, fort.' The first word I say will always be found in the first column of your scratch out answer sheet, and will be one of these four words, (indicating by pointing to the group of four words from which the first word is to be chosen), the second word will always be one of these four words (indicating by pointing to the group of four words from which the secone is to be chosen) and the last word will always be one of the four words in the last column. You are to listen to all three words in each series, and then scratch out the three words you think I said. Remember, the first word will be in the first column, the second word in the second column, and the third word I say will be in the last column. When we finish with series number 1, we will go to series number 2 and then all the way down the list through series number 9.
>
> Sometimes the words will sound so soft you may not hear them well, and at other times a list may sound comfortably loud to you. Sometimes you will be able to hear me and to see me, and sometimes I will allow you only to hear me. If I tell you that you are to look and listen, you are to look at me while I say the words. If I tell you that you are to listen only, you are to turn in your chair and look at this wall (pointing to the wall to the subject's right).
>
> Before saying a word list I will tell you three things: First, the type of list we will use, that is, whether it will be a 'Say the Word List' or a 'Scratch Out List'; Secondly, whether you can listen only, or can look and listen; and Thirdly, I will tell you whether it will be louder, softer, or about as loud as the previous list presented.

If the subject was unsure of what to do, the procedure was again explained, and more examples were given. The subject was then told in which ear he would hear the experimenter, and the headphones were placed over the ears. The experimenter then entered the control room and the speech reception threshold was obtained.

Speech thresholds were obtained for only the ear to be used in the word recognition tests, which was the same ear that met the subject selection criteria established for a group.

For the normal hearing subjects, the first spondaic word was presented at an intensity level approximately 15 db above threshold for speech (44 db re .0002 microbar). With each correct response the intensity of speech was reduced 2 db until the subject missed a word. At this level, six to ten spondaic words were presented. If the subject failed to recognize correctly 50 per cent of the words, the intensity level was increased 4 db, and 6 to 10 more words were given. The intensity was then decreased in 1 db steps and 6 to 10 words given at each level until a word recognition score of 50 per cent was obtained.

For the hard of hearing subjects, speech reception thresholds were established in essentially the same manner as for the normal hearing subjects. The single exception was that the first spondaic word was presented at a sensation level approximately 15 db above the subject's average hearing loss for pure tones in the speech range, rather than at a level of approximately 15 db above the normal threshold for speech.

The speech reception threshold was then used as the reference point for measuring intensity levels for the presentation of the word lists under experimental conditions.

Experimental Testing Procedures

Immediately after establishing the subject's speech reception threshold, testing under the experimental conditions began. Before a list of words was read, the subject was given the following information: First, the subject was informed of the type of word list to be communicated, that is, whether it would be a PB List or a Black List. If a Black List was to be given, the list number on the experimenter's testing schedule was checked with the answer sheet in front of the subject. This was to make certain that both the subject and the experimenter had the lists in the same order. Secondly, the subject was informed of the condition of presentation, that is, whether it would be an auditory, visual, or bisensory test. If the testing schedule called for an auditory presentation, the subject was instructed to turn in his chair and face the wall of the test room. This prevented him from observing the experimenter through the window which connected the two rooms. Under the auditory condition, then, no visual cues of speech were available to the listener.

If the testing schedule indicated that a bisensory presentation was to be made, the subject was instructed to face the experimenter and to watch him as well as to listen to him as he said the words. Under the bisensory condition, both the visual and auditory cues of speech were available to the listener.

Finally, when the purely visual presentation was to be made, the subject was required to take off the earphones and to watch the experimenter through the window as the word list was presented. In addition, the microphone feeding the speech audiometer was turned off. Therefore, under the visual condition, no auditory cues of speech were available to the subject.

The final bit of information given the subject pertained to the relative intensity level of the presentation. Each subject was told whether the presentation of a word list would be at a level softer, louder, or about as loud as the previous presentation. In the case of the first experimental list, the spondaic words presented to establish the speech reception threshold were considered to be the previous presentation.

Once the information regarding the type of list, condition of presentation, and relative intensity level was given the subject, the attenuator was set for presentation of the speech material. Speech samples were presented through a single channel of the speech audiometer, and instructions were presented through a second channel at a sensation level which provided easy and accurate communication with the subject.

A carrier phrase or monitoring phrase, "Say the Word ________," was used when presenting the phonetically balanced lists. The monitoring phrase and stimulus word were presented by the experimenter and the subject merely repeated the word. If the response was incorrect, the experimenter penciled through the stimulus word appearing on the word list before him. When a Black List was presented, the experimenter stated the number of the series, and then presented the sequence of three stimulus words. The subject was given ample time to pencil through the three words he selected from the multiple choice list before him. After all word lists were presented, each list was scored for correct responses. The raw scores were then recorded on a record sheet prepared for each subject.

Summary

The purpose of the study was to evaluate the role of visual and auditory cues in verbal communication. To make this evaluation, three groups of subjects were used. One group contained subjects with normal hearing. A second group was composed of subjects with mild hearing losses and relatively flat audiometric configurations, and the third group was made up of individuals who had mild hearing losses with sloping configurations.

Testing was conducted in the sound-proofed rooms of the Speech and Hearing Clinic of the University of Oklahoma, with the aid of a commercial pure tone audiometer, and a custom built partially transistorized speech audiometer. Phonetically balanced word lists and multiple-choice word lists were used as speech materials.

All subjects received a preliminary air conduction and bone conduction test of hearing acuity, and instructions were then given regarding the experimental testing procedures. Each type of word list was presented under three experimental conditions: auditory, visual and bisensory. Under the auditory condition, the subject was presented with auditory cues only, while under the visual condition, only visual cues were available to the subject. Finally, under the bisensory condition, the subject was presented with both the visual and auditory stimuli. The auditory and bisensory presentations of speech samples were made through the speech audiometer at eight sensation levels, ranging from -6 db through 15 db in 3-db steps. Only one visual presentation was made for each list type. Using this procedure, word recognition scores were obtained for an auditory, a visual and a bisensory presentation of speech.

CHAPTER IV

RESULTS

Introduction

In this study, word recognition scores (number of words correctly recognized) were obtained for three groups of subjects. Group I consisted of subjects with normal hearing, Group II subjects with flat hearing losses, and Group III subjects with sloping hearing losses. Each subject responded to two types of word lists (the PB Lists and the Black Lists) under three different experimental conditions. Under the first, only the auditory cues of speech were available to the listener, while under the second, word recognition scores were obtained when both the auditory and the visual cues were available. Scores for the auditory and bisensory presentations (see Table 1) were obtained at eight sensation levels ranging from -6 db through 15 db. A discussion of the results reported below will appear in the following chapter.

Results Obtained with the Phonetically Balanced Word Lists

Auditory, Bisensory and Visual Scores

In Figure 2 are plotted the mean auditory, bisensory and visual recognition scores for the PB Lists for Groups I, II and III. Since sensation level does not affect the visual score, and since only one visual presentation was made for each list type, the visual score is shown as a straight line.

Auditory recognition scores.--The minimum mean auditory recognition scores for Groups I, II and III were similar, ranging only from 7.5% to 8.2%. The maximum mean auditory scores for Groups I and II were identical (85%), while the highest mean score for Group III was somewhat lower (64.5%). Group III not only differed in terms of the maximum intelligibility scores, but inspection of Figure 2 will also reveal that for Group III the intelligibility of words increased less rapidly as a function of intensity than it did for Groups I and II. It can be seen that the divergence of the scores of Group III from those of Groups I and II becomes apparent at the -3 db sensation level.

Visual recognition scores.--The visual recognition scores earned with the PB Lists are also shown in Figure 2. It is readily seen that Group III earned the lowest word intelligibility score of the three groups when only the visual cues of speech were available. The score of 5.3% for Group III can be compared to scores of 11% for Group I and 9.3% for Group II. A t-test of differences between means of the groups indicated that a difference of statistical

TABLE 1

MEAN NUMBER AND PER CENT OF WORDS FROM PB LISTS REPRODUCED CORRECTLY BY THREE GROUPS OF SUBJECTS UNDER THREE CONDITIONS OF PRESENTATION (VISUAL, AUDITORY, BISENSORY)

Group	Sensation Level of Auditory Stimulus in DB	Visual No.	Visual %	Auditory No.	Auditory %	Bisensory No.	Bisensory %
		5.5	11.0				
	-6			4.0	8.0	14.4	28.8
	-3			8.8	17.5	21.1	42.3
	0			12.9	25.8	28.8	57.5
I	3			19.4	38.8	33.2	66.5
	6			29.6	59.2	37.2	74.5
	9			35.0	70.0	42.8	85.5
	12			38.1	76.2	44.8	89.5
	15			42.5	85.0	47.2	94.5
		4.6	9.3				
	-6			4.1	8.2	15.0	30.0
	-3			7.1	14.2	21.1	42.2
	0			14.8	29.5	28.8	57.5
II	3			22.0	44.0	35.6	71.2
	6			30.9	61.8	40.1	80.2
	9			34.4	68.8	42.6	85.2
	12			40.2	80.5	44.8	89.5
	15			42.5	85.0	47.2	94.5
		2.6	5.3				
	-6			3.8	7.5	14.8	29.5
	-3			6.5	13.0	22.1	44.2
	0			9.1	18.2	26.8	53.5
III	3			13.6	27.2	30.9	61.8
	6			19.1	38.2	35.5	71.0
	9			24.8	49.5	40.1	80.2
	12			31.2	62.5	41.5	83.0
	15			32.2	64.5	44.5	89.0

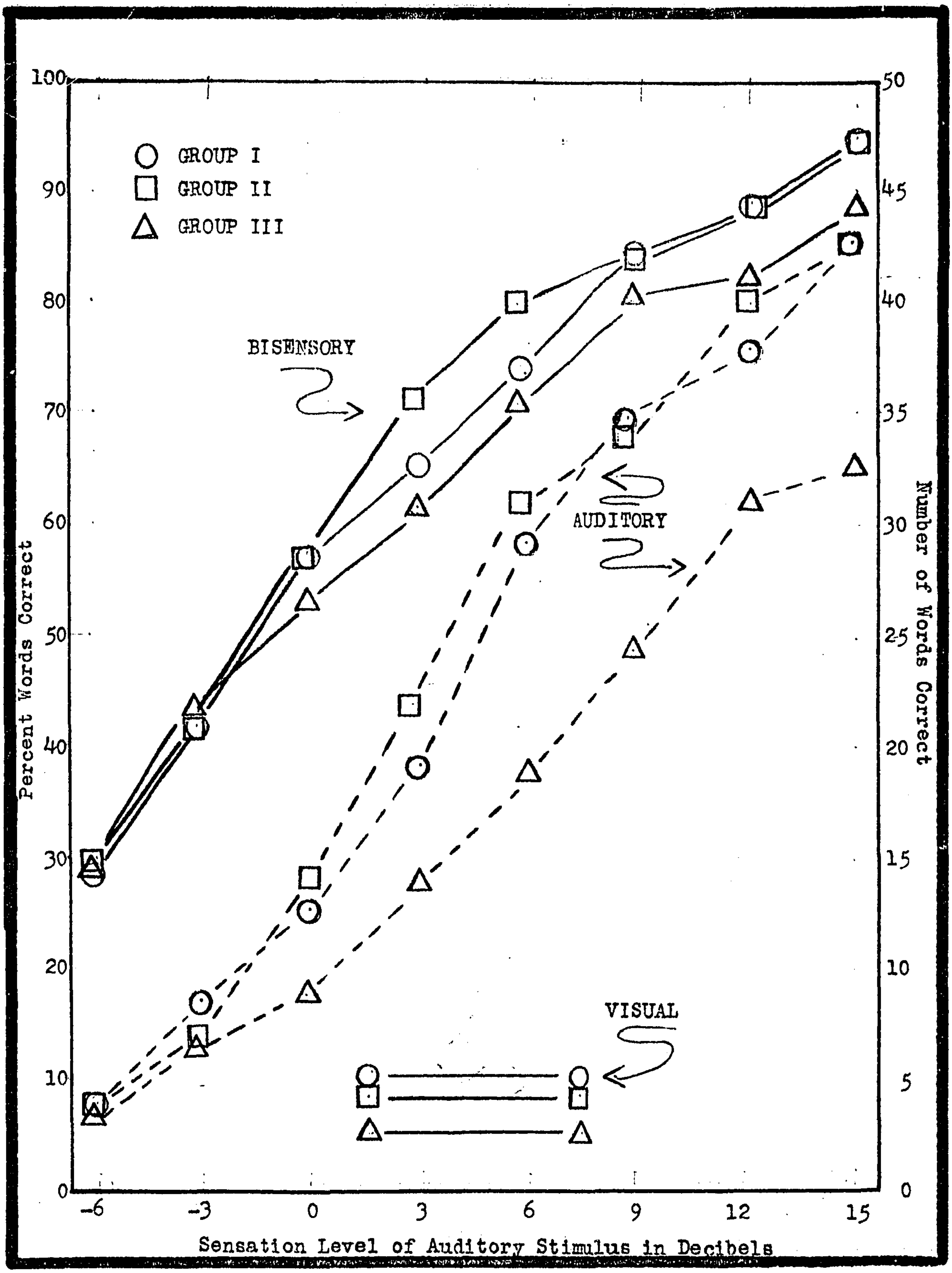

Figure 2. Auditory, visual and bisensory word recognition scores for three groups of subjects when presented with PB word lists.

significance (.02   .05) occurred only between the means of Groups I and III.

Bisensory recognition scores.--The minimum mean bisensory intelligibility scores were similar for the three groups, having values of 28.8% for Group I, 30.0% for Group II, and 29.5% for Group III. The maximum mean bisensory intelligibility scores for Groups I and II were the same (94.5%), while a lower score of 89.0% was earned by Group III. For the bisensory presentations, as in the case of the auditory presentations, the intelligibility of words increased more rapidly as a function of intensity for Groups I and II than it did for Group III. Here, too, it can be observed that the deviation in the performance of Group III from that of Groups I and II begins at the 0 db sensation level.

When the auditory and bisensory scores were compared, certain features of group performance became evident. It is especially important to point out that Group III earned considerably poorer auditory scores than Groups I and II, particularly at sensation levels above 3 db, although the bisensory scores of the three groups did not deviate from one another to as great an extent as did the auditory scores. It appears, then, that Group III was able to make greater improvement in the intelligibility of speech owing to the use made of the combined visual and auditory cues than were Groups I and II.

A clearer picture of differences in group performance can be made by comparing the auditory reception of speech to the bisensory reception of speech at each sensation level. The degree of improvement provided by the inclusion of visual cues was assessed by subtracting the auditory score at a given sensation level from the bisensory score of the same level. The difference score thus obtained was interpreted as the relative amount of improvement in intelligibility over the auditory presentation as a result of having made visual cues available to the listener.

The difference between the auditory scores and the bisensory scores (difference scores) are plotted in Figure 3. The difference scores between the auditory and bisensory presentations are similar for Groups I and II. For Group III, however, greater differences occurred between the auditory and bisensory scores than for the other two groups. A feature common to all three groups, however, is the curvilinearity of the parameter representing the difference in intelligibility of speech between the auditory and bisensory presentations. As anticipated, the greatest differences were present at or near auditory threshold levels.

Not only did the auditory and bisensory performances vary as a function of the sensation level of the auditory pattern, but there were also group differences. For Group I, the least amount of improvement in intelligibility was approximately 9% at the 15 db sensation level. For Group II, the least improvement in intelligibility of speech also amounted to approximately 9% at the 12 db and 15 db sensation levels. And for Group III, the least amount of improvement was 20% at the 12 db sensation level.

The maximum improvement in intelligibility for Group I was 32% and for Group II, 28%. For Group III, the greatest improvement in intelligibility (bisensory vs. auditory reception of speech) was 36%. For all three groups, the greatest differences between the auditory and bisensory scores occurred at the zero decibel sensation level.

It becomes apparent from these results that the expected interaction between the auditory and visual cues of speech provided for a varying degree of improvement over the isolated auditory or visual reception of speech, depending upon the intensity level of the auditory stimulus. The varying effect of the interaction of the visual and auditory components of speech makes it impossible to assign an absolute quantitative value to the contribution of either of the two components to a bisensory reception of speech. Thus, the difference scores shown in Figure 3 may be taken as the effects of the interaction between visual and auditory cues. It is this interaction which provided for the maximum understanding of speech at a given sensation level.

It has been pointed out, then, that the bisensory word recognition scores were better than the auditory scores of the same sensation level, and that improvement in the

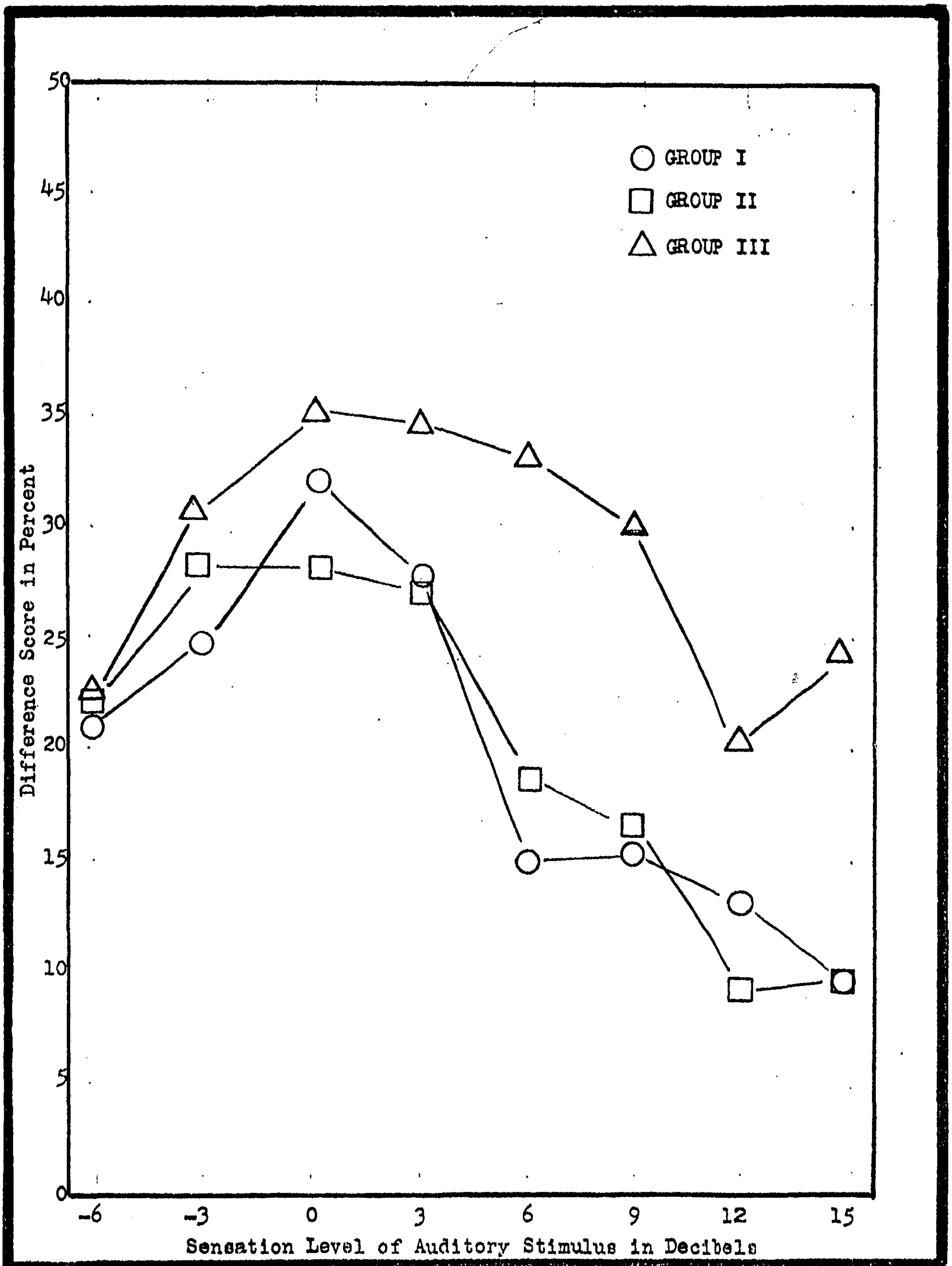

Figure 3. Difference scores (bisensory score minus auditory score) for three groups of subjects when presented with PB word lists.

intelligibility of a bisensory presentation over an auditory presentation of speech varied depending upon the sensation level of the auditory component. However, the effects of the interaction of vision and audition in the reception of speech can also be looked at from another point of view. It is possible to compare the bisensory word recognition scores with the sum of the visual and the auditory scores. Such a comparison can be made on the hypothesis that if no interaction took place between the two components under study, the bisensory scores at a given sensation level should approximate the sum of the visual score and the auditory score at the same sensation level.

Figure 4 is a graphic representation of the differences between the sum of the auditory and visual scores and the bisensory score. The parameters in the figure were derived by subtracting the arithmetic sum of the auditory and visual scores from the bisensory discrimination score. Using this procedure, a positive value indicates that the bisensory score exceeded the arithmetic sum of the auditory and visual scores, while conversely, a negative value indicates the opposite to be true. Figure 4 shows that the bisensory scores (with the exception of Group I at the 15 db level and Group II at the 12 db level) are always greater than the sum of the visual and auditory scores. When the PB Lists are used as the speech samples, the general effect of the interaction between the visual and auditory components

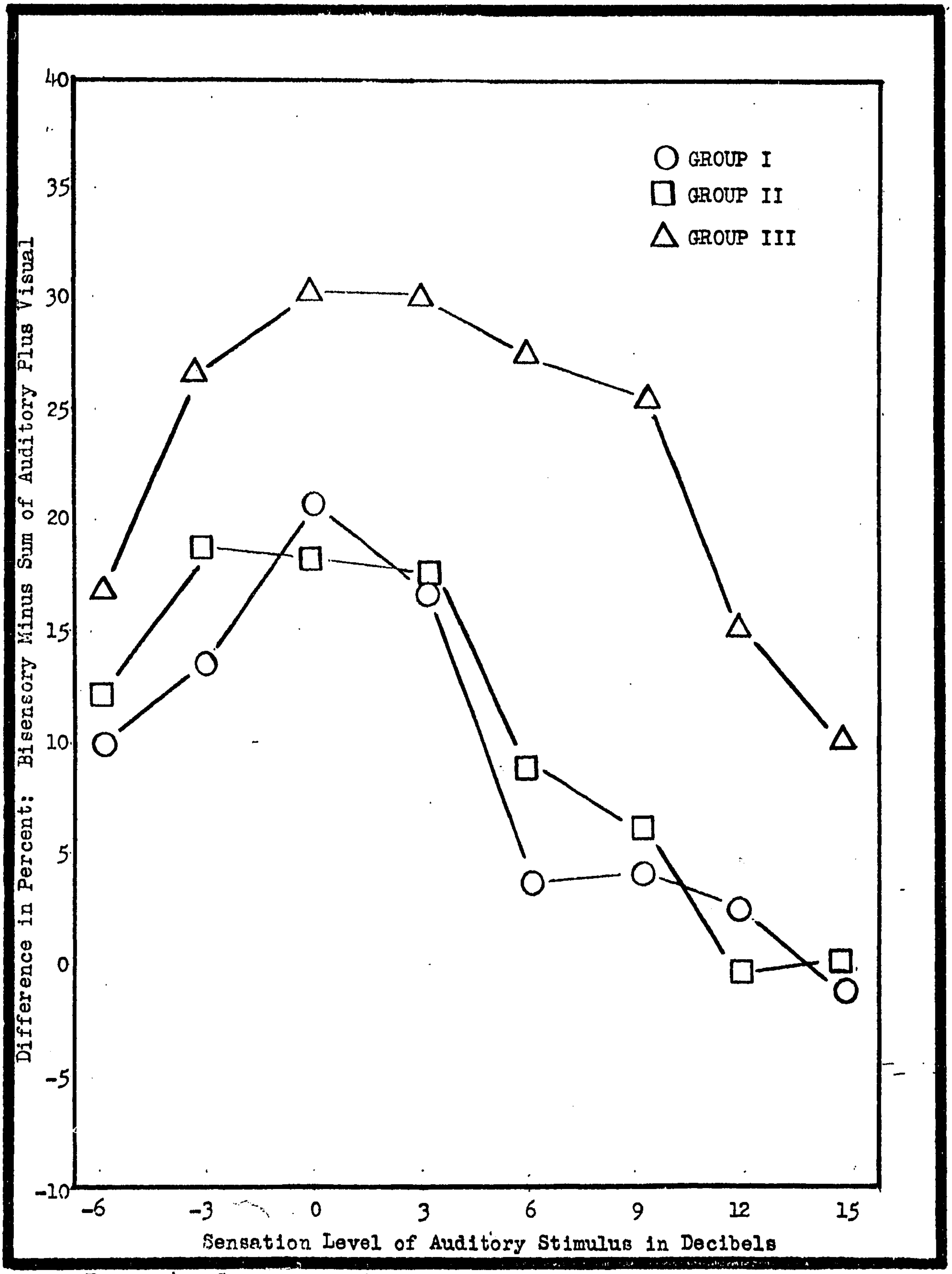

Figure 4. Comparison of bisensory performance with arithmetic sum of auditory and visual performances of three groups presented with PB word lists.

is to improve the intelligibility of speech beyond the value derived by simply summing the visual and auditory discrimination scores.

Statistical Analysis

In an effort to explore the differences among the various word recognition scores, the data were analyzed with an analysis of variance technique. The first design used was a complex mixed factorial design containing the variables of (1) conditions (auditory and bisensory), (2) intensity (the eight sensation levels), (3) groups (three groups) and (4) word lists (PB Lists and Black Lists). In the analysis, the main effects of each of the variables were not clear because of the first and second order interactions among the variables. Table I is a summary of this complex analysis. The results indicated that no statement regarding the interaction of visual and auditory cues could be made without first considering the intensity level of the auditory cues, the hearing acuity of the subjects, and the type of speech material presented to the subjects. This finding, based upon a statistical analysis, was to be expected after reviewing the graphs of the data.

It was necessary to re-evaluate the data in an effort to study in a more refined manner the effects of combined visual and auditory cues of speech on intelligibility of oral communication. In this analysis, each of the word lists was considered separately. In addition, each sensation level was considered separately for each word list. Thus, an analysis of variance with a model suggested by Edwards,[1] was made at each sensation level for each word list, the variables being: (1) conditions and (2) groups. Table 3 is a summary of the findings revealed by the analysis of variance for groups and for conditions when subjects were presented with the PB Lists.

Inspection of Table 3 points out that statistically significant differences between groups occur at sensation levels of 3 db, 6 db, 9 db, 12 db and 15 db. Referring to Figure 3, it can be seen graphically that the deviation of difference scores of Group III from those of Groups I and II is most apparent at the sensation levels of 3 db and above. It can also be observed (Figure 2) that beginning at the 3 db sensation level, both the auditory and the bisensory recognition scores were consistently poorer for Group III than for Groups I and II. The poorer auditory and bisensory scores for Group III may account for the statistically significant difference among group means at the 3 db, 6 db, 9 db, 12 db and 15 db sensation levels.

Inspection of Table 3 also shows that for every sensation level tested, the differences between the auditory recognition scores and the bisensory recognition scores were statistically significant.

---

[1]A. L. Edwards, Experimental Design in Psychological Research (New York: Rinehart and Co. Inc., 1950), Chapter XV.

TABLE 2

A SUMMARY OF THE ANALYSIS OF VARIANCE FOR CONDITIONS, WORD LISTS, INTENSITY LEVELS AND GROUPS

Source	Sum of Squares	Mean Squares	Degrees of Freedom	F. Value	Significance
Group	1694.0	847.00	2	2.19	N.S.
Subjects within groups	8113.0	386.33	21		
Intensity	47498.0	6785.43	7	303.05	**
I X G	644.0	46.00	14	2.05	*
I X S	3291.0	22.39	147		
Lists	20481.0	20481.00	1	242.87	*
L X G	249.0	124.50	2	2.48	N.S.
L X S	1771.0	84.33	21		
L X I	9971.0	1424.43	7	105.50	**
L X I X G	219.0	15.64	14	1.12	N.S.
L X I X S	2004.0	13.63	147		
Conditions	12610.0	12610.00	1	229.06	**
C X G	227.0	113.50	2	2.06	N.S.
C X S	1156.0	55.05	21		
C X I	1002.0	143.14	7	13.69	**
C X I X G	220.0	15.71	14	1.50	N.S.
C X I X S	1536.0	10.45	147		
C X L	2120.0	2120.00	1	99.81	**
C X L X G	248.0	124.00	2	5.84	**
C X L X S	446.0	21.24	21	2.54	*
C X L X I	310.0	44.29	7	5.30	**
C X L X I X G	117.0	8.36	14		
C X L X I X S	1262.0	8.59	147		
Total	117216.0	152.82	767		

N.S. Not Significant
* Significant beyond 5% level
** Significant beyond 1% level

TABLE 3

SUMMARY OF THE FINDINGS OF THE ANALYSIS OF VARIANCE FOR GROUPS AND CONDITIONS FOR THE PB LISTS

Sensation Level	-6 db	-3 db	0 db	3 db	6 db	9 db	12 db	15 db
Groups I, II, and III)	N.S.	N.S.	N.S.	*	*	*	*	**
Conditions (Auditory and bisensory)	**	**	**	**	**	**	**	**
Interaction: Groups X Conditions	N.S.	N.S.	N.S.	N.S.	**	*	N.S.	**

N.S. Not Significant
* Significant beyond the 5% level
** Significant beyond the 1% level

Finally, a trend toward interaction of groups and conditions is observed at the sensation levels of 6 db and above. Referring now to Figure 3, the trend toward interaction can be seen graphically. The parameter representing difference scores for Group III failed to parallel the parameters for Groups I and II. This lack of parallelism of the parameters is a graphic indication of an interaction between groups and conditions. The interaction of groups and conditions, which reaches statistical significance at the 6 db, 9 db and 15 db sensation levels, may be interpreted to mean that at these levels the effects of conditions on the performance of the groups was <u>similar</u> at the sensation levels of 3 db and below, but that a trend toward <u>differences</u> in performance was apparent at levels of 6 db and above. It means, further, that in discussing the effects of conditions, it is necessary to know the intensity level of the auditory cues and, in addition, to identify the group of subjects under consideration.

Results Obtained with the Black Multiple-Choice Lists

Auditory, Visual and Bisensory Scores

In Table 4 are reported the auditory, visual and bisensory recognition scores for the Black Lists. As in the case of the PB Lists, the bisensory scores for the Black Lists were appreciably better than the auditory scores. In general, differences in intelligibility between the auditory and the bisensory reception of speech tended to become less and less as the intensity of speech was increased. The trend was less apparent, however, for Group III than for Groups I and II.

<u>Auditory recognition scores</u>.--The minimum mean auditory recognition scores (Figure 5) were 20.8% for Group I and 19.0% for Group II, as compared with 31.0% for Group III. The maximum mean auditory recognition scores were 83.3% for Group I, 89.8% for Group II and 68.5% for Group III. With the Black Lists, as in the case of the PB Lists, subjects

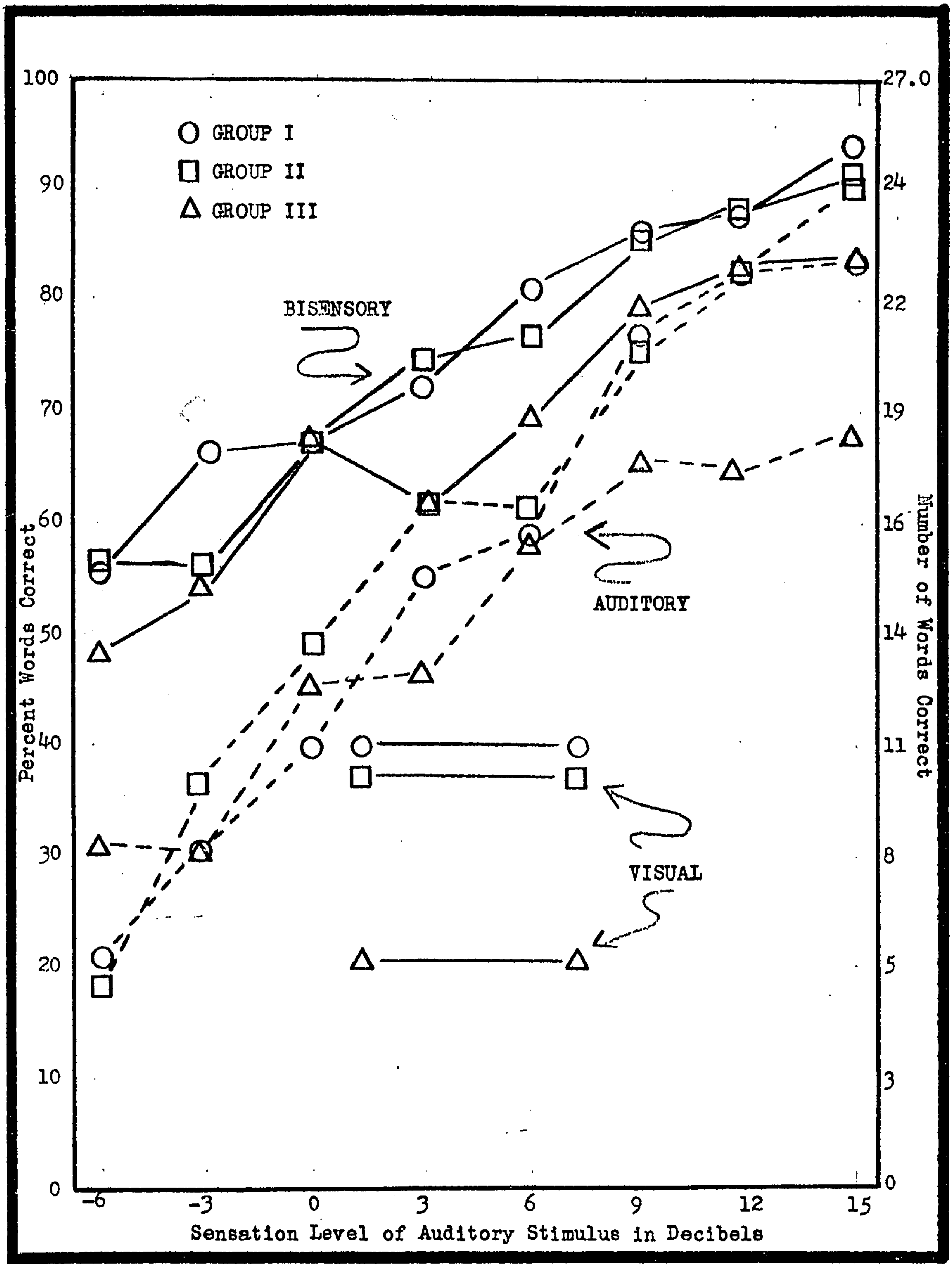

Figure 5. Auditory, Visual and bisensory word recognition scores for three groups of subjects when presented with multiple-choice word tests.

in Group III failed to attain maximum auditory scores comparable to those of Groups I and II. It was also noted that the increase in word intelligibility as a function of intensity was not as rapid for Group III as it was for the other two groups.

When the auditory reception of speech is considered, it is seen that all three groups earned higher word intelligibility scores for the Black Lists than for the PB Lists, with the better intelligibility for the Black Lists most apparent at the sensation levels of -6 through 9 db. For instance, at the -6 db sensation level, Group I earned an intelligibility score of only 8.0% with the PB Lists as compared to 20.8% with the Black Lists. Group II had a score of 8.2% with the PB Lists and 19.0% with the Black Lists. Finally, Group III obtained an intelligibility score of 7.5% with the PB Lists, as compared to 31.0% with the Black Lists. At the 15 db sensation level, where intelligibility scores earned with each list were similar, Group I earned a score of 85.0% with the PB Lists compared to 83.3% for the Black Lists. Group II also earned a score of 85.0% with the PB Lists, but 89.8% with the Black Lists. Group III, with the lowest auditory intelligibility scores, obtained 64.5% correct with the PB Lists and 68.5% with the Black Lists. In general, when auditory presentations are made to a given group of subjects, the differences in intelligibility between the PB Lists and the Black Lists are greater at the

TABLE 4

MEAN NUMBER AND PER CENT OF WORDS FROM BLACK LISTS REPRODUCED CORRECTLY BY THREE GROUPS OF SUBJECTS UNDER THREE CONDITIONS OF PRESENTATION (VISUAL, AUDITORY, BISENSORY)

Group	Sensation Level of Auditory Stimulus in DB	Visual No.	Visual %	Auditory No.	Auditory %	Bisensory No.	Bisensory %
		10.9	40.3				
	-6			5.6	20.8	15.0	55.6
	-3			8.2	30.6	17.9	66.2
	0			10.8	39.8	18.1	67.1
I	3			15.0	55.6	19.6	72.7
	6			16.0	59.3	21.9	81.0
	9			20.5	75.9	23.4	86.6
	12			22.1	81.9	23.6	87.5
	15			22.5	83.3	25.4	94.0
		9.9	36.6				
	-6			5.1	19.0	15.4	57.0
	-3			9.9	36.6	15.2	56.5
	0			13.4	49.6	18.1	67.1
II	3			16.8	62.0	20.1	74.5
	6			16.6	61.6	20.5	75.9
	9			20.1	74.5	23.2	86.1
	12			22.1	81.9	24.0	88.9
	15			24.2	89.8	24.6	91.2
		5.5	20.4				
	-6			8.4	31.0	13.0	48.2
	-3			8.2	30.6	14.6	54.2
	0			12.2	45.4	18.2	67.6
III	3			12.5	46.3	16.8	62.0
	6			15.8	58.3	18.8	69.4
	9			17.8	65.7	21.4	79.2
	12			17.4	64.4	22.0	81.5
	15			18.5	68.5	22.4	82.9

lower sensation levels (-6 db and -3 db) than at the higher sensation levels (12 db and 15 db).

<u>Visual recognition scores</u>.--The mean visual recognition scores for the three groups of subjects when presented with the Black Lists are also shown in Figure 5. Inspection of the figure points out that with the Black Lists (as in the case of the PB Lists), Group III earned the lowest mean visual score. In terms of percentage of words correct, Group I earned a score of 40.3%, Group II 35.5% and Group III only 20.4%. In view of these large differences among the mean scores of the three groups a t-test of the difference between means was computed. The results of the t-test indicated that a statistically significant difference exists between the means of Groups I and III (.001   .01) and Groups II and III (.02   .05).

When comparing the visual recognition scores obtained with the Black Lists to those recorded for the PB Lists, certain features can be noted. For one thing, the intelligibility scores for the Black Lists were higher for all three groups than the scores for the PB Lists. For the PB Lists, scores of 11.0%, 9.3% and 5.3% were earned by Groups I, II and III respectively, while with the Black Lists, intelligibility scores of 40.3%, 36.6% and 20.4% were recorded for Groups I, II and III.

A second feature worthy of note was that although the intelligibility scores of the groups varied from one type of word list to another, the relative position of the group scores did not change. In each instance, Group III obtained the lowest visual recognition scores and Group I the highest, leaving the performance of Group II somewhere between that of Groups I and III.

Bisensory recognition scores.--When the Black Lists were presented with bisensory stimulation, Group I earned a minimum mean score of 55.6% and Group II a score of approximately the same value, 57%. However, Group III obtained a somewhat poorer minimum score of 48.2%. The maximum mean scores for the bisensory presentations were 94% for Group I, 91.2% for Group II and 82.9% for Group III.

At sensation levels of -6 db through 0 db, the intelligibility scores of the three groups obtained from the bisensory presentations of the Black Lists were higher than those of the PB Lists. Inspection of Figure 2 shows that at the -6 db sensation level, the bisensory intelligibility scores for the three groups ranged from 28% to 30% when the PB Lists were used, as compared to the performance on the Black Lists (see Figure 5), where the lowest bisensory scores ranged from 48% to 55% for the same three groups. This trend continued through the 3 db sensation level. However, at sensation levels above 3 db, the intelligibility scores obtained with a bisensory presentation of the Black Lists were either comparable to, or slightly less, than those obtained with the PB Lists.

With the Black Lists, as with the PB Lists, a comparison of the bisensory reception of oral communication with the auditory reception was made. Figure 6 shows the difference scores (bisensory scores minus auditory scores) for Groups I, II and III. Here it can be seen that when the Black Lists were used, there was a rather progressive decrease in differences between the bisensory and auditory performances as the intensity of speech increased. The difference scores obtained with the Black Lists differed from those of the PB Lists in two respects. First, the curvilinearity of the parameter (difference scores) observed when the PB Lists were used (see Figure 3) was not seen for the Black Lists (see Figure 6). Secondly, the devious performance of Group III, in terms of larger differences between the bisensory and auditory presentations of the PB Lists, was not apparent when the Black Lists were employed.

An interesting contrast appears in the relationship of the auditory, visual and bisensory scores for the Black Lists when compared with the relationship of the same word recognition scores obtained with the PB Lists. It will be recalled that for the PB Lists, the bisensory scores for all three groups were, in general, <u>higher</u> than the sum which resulted from simply adding the visual scores and the auditory score. With the Black Lists, however, the bisensory recognition scores were, in general, <u>lower</u> than the sum of the auditory score plus the visual score. Figure 7 shows the

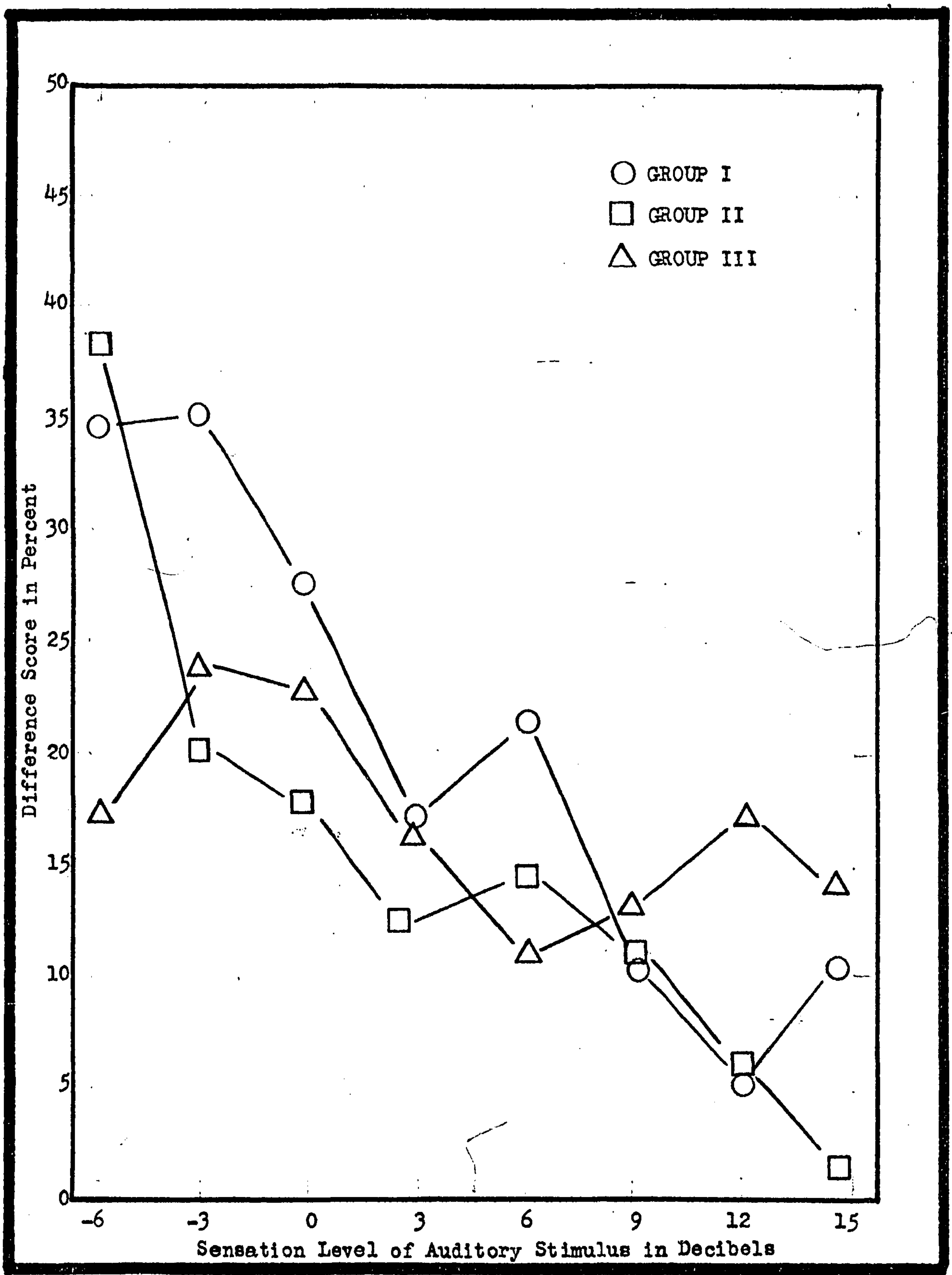

Figure 6. Difference scores (bisensory score minus auditory score) for three groups of subjects when presented with multiple-choice word tests.

differences between the bisensory performance and the sum of the visual and auditory scores for the Black Lists. It seems, then, that when multiple-choice word lists like the Black Lists are employed, the effects of the interaction of the visual and auditory cues of speech differ from the effects when the PB List are used (see Figure 4). Affecting the interaction is the fact that the visual and auditory scores were higher for the Black Lists than for the PB Lists, and therefore, the sum of the visual and auditory scores more often exceeded the bisensory scores. One feature, however, remains similar, namely that the deviation in performance of Group III from that of the other two groups is observed for both list types.

Statistical Analysis

As with the PB Lists, differences among groups and between conditions of presentation of the Black Lists were evaluated at each intensity level, utilizing an analysis of variance technique. Table 5 summarizes the findings for each intensity level.

As can be seen, the mean auditory recognition scores differed significantly from the mean bisensory recognition scores at every sensation level tested. Thus, when speech was presented at sensation levels of from -6 db to +15 db, the supplemental visual cues resulted in an appreciable increase in the word recognition scores. When considering groups, however, a statistically significant difference

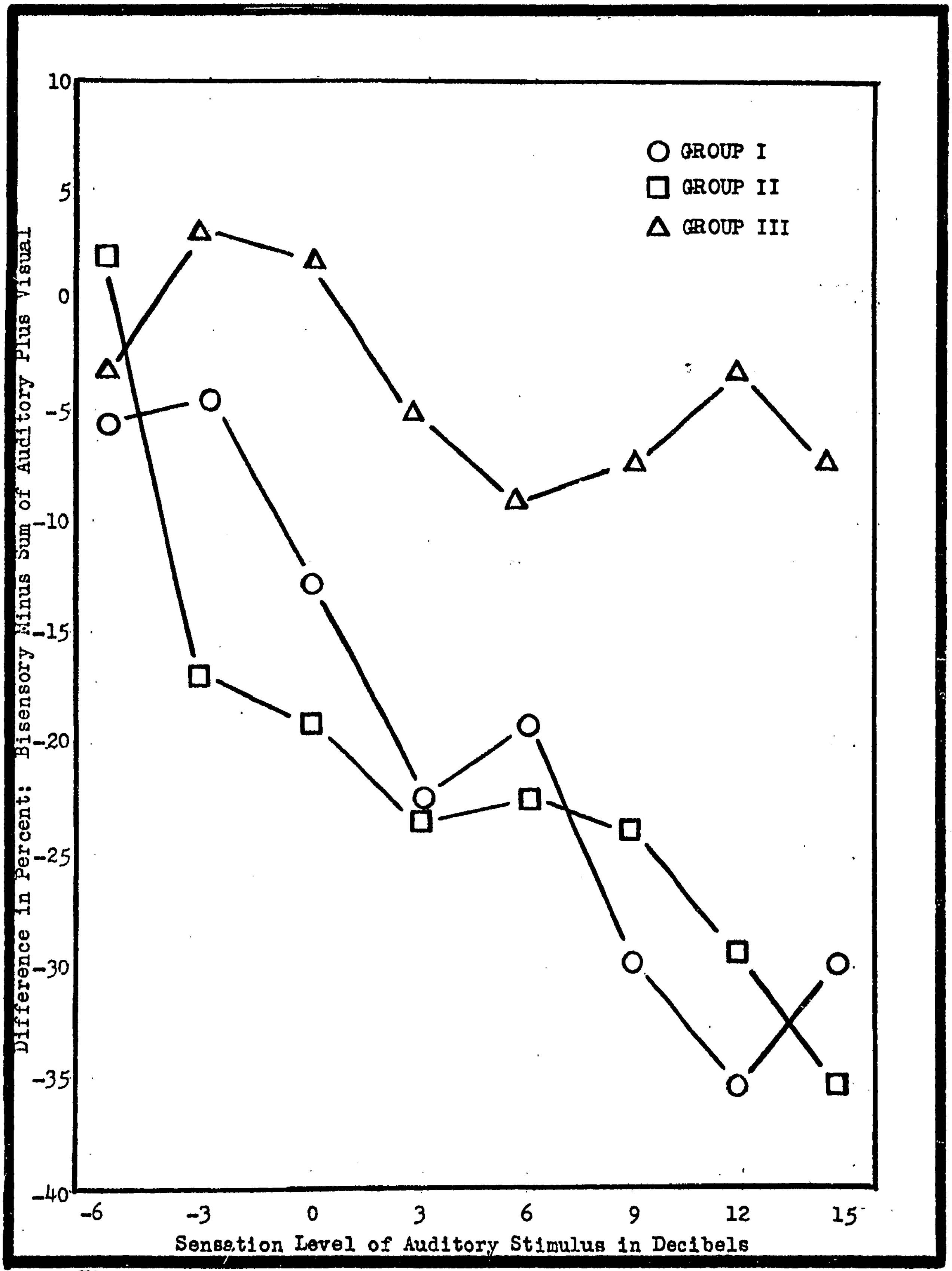

Figure 7. Comparison of bisensory performance with arithmetic sum of auditory and visual performances of three groups presented with multiple-choice word lists.

among groups occurred only at the 15 db sensation level. The difference in performance between groups at the 15 db level may have occurred as a result of the scores of Group III deviating from those of Groups I and II. It can be seen in Figure 5 that for Groups I and II, the auditory scores continue to increase in magnitude through the 15 db sensation level. However, for Group III, the increase is not as evident at sensation levels above 9 db.

TABLE 5

SUMMARY OF THE FINDINGS OF THE ANALYSIS OF VARIANCE FOR GROUPS AND CONDITIONS FOR THE BLACK LISTS

Sensation Level	-6 db	-3 db	0 db	3 db	6 db	9 db	12 db	15 db
Groups I, II, and III	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.	**
Conditions (Auditory vs. bisensory)	**	**	**	**	**	**	*	**
Interaction: Groups X Conditions	*	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.

N.S. Not Significant
* Significant beyond the 5% level
** Significant beyond the 1% level

At only one sensation level did the interaction between conditions and groups prove to be of statistical significance. The interaction occurred at a sensation level of -6 db, and was apparently the result of Group III obtaining a considerably better auditory score than Groups I and II at this one level. At all other sensation levels, Group III failed to earn auditory recognition scores equal to, or higher than, those of Groups I and II (see Figure 3).

Summary

Performance scores for three groups of eight subjects each were obtained with two types of word lists at eight sensation levels. Words were presented under three conditions: auditory, visual and bisensory. The results can be summarized in the following manner:

1. The bisensory reception of speech allows greater intelligibility than does an auditory or visual reception of speech. This superiority of the bisensory reception holds true for all three groups of subjects and from intensity level to intensity level of the auditory component.
2. The understanding of speech was affected by the intensity level of the auditory cues in the following manner:
    a. When the PB Lists were used, the greatest degree of improvement in word intelligibility resulting from the interaction of visual and auditory cues occurred at threshold levels. This was true for all three groups of subjects.
    b. For the Black Lists, however, the greatest

improvement in speech intelligibility resulting from the interaction of the visual and auditory cues occurred at the -6 db sensation level for Group I and II, and at the -3 db sensation level for Group III.

c. In general, the differences in intelligibility of speech between its auditory and bisensory reception became less marked as the intensity of the auditory component increased.

3. The lowest mean auditory recognition scores for a given word list were of similar magnitude from group to group. However, the maximum auditory scores differed in magnitude, with Group III earning scores falling short of those earned by Groups I and II.
4. The lowest mean bisensory recognition scores when using a given word list were comparable in magnitude from group to group. In contrast, the highest bisensory scores of Group III fell short of those earned by Groups I and II.
5. The deviation in performance of Group III from that of Groups I and II, in terms of word recognition scores, was more pronounced for the auditory scores than for the bisensory scores. This was true for both types of word lists.
6. With both word lists, Group III made a greater im-

provement in word intelligibility as a result of the interaction of visual and auditory cues than did Groups I and II.

7. The bisensory recognition scores cannot be considered as the simple summation of the auditory and the visual recognition scores. For the PB Lists, the bisensory scores were greater than the sum of the auditory and visual scores. Conversely, the bisensory scores for the Black Lists were smaller than the sum of the visual and the auditory recognition scores.

CHAPTER V

DISCUSSION

Introduction

In this chapter the results of the study, reported in Chapter IV, will be compared with those of other investigators. In addition, some of the factors which contributed to the findings will be discussed and finally, some of the implications of these findings will be considered.

Auditory Recognition Scores

The speech intelligibility scores for Groups I and II obtained with the PB Lists are in keeping with those reported by Egan.[1] A relatively linear increase in word recognition scores occurred as a result of increases in the intensity of speech (within the sensation levels tested). Quantitatively, the increases amounted to approximately 12% per 3 db intensity gain when a line of best fit was applied to the auditory articulation function.

Group III (sloping loss) obtained poorer auditory

---

[1] J. P. Egan, "Articulation Testing Methods II," Laryngoscope LXII (1952), pp. 955-991.

reccgnition scores with the PB Lists than did Groups I (normal hearing) and II (flat loss). These findings corroborate those of Davis,[1] who stated that: "...a man with a severe high-tone nerve deafness will always fail to hear certain sounds and never make a perfect articulation score." The relative inability to hear the higher frequency sounds of speech when presented at a given sensation level is felt to account for the poorer auditory recognition scores of Group III. The increase in word intelligibility of approximately 8% per 3 db gain in intensity for Group III can be compared to the 12% increase per 3 db gain in intensity for Groups I and II. It appears, then, that the relative inability to hear high frequency sounds of speech affected not only the maximum auditory intelligibility scores but also the slope of the articulation function.

The relatively slower rise in the articulation function of Group III (see Figure 2) is important for two reasons. First, it could account for the statistically significant differences among groups which are apparent at the sensation levels of 6 db, 9 db, 12 db and 15 db (see Table III). Secondly, the more slowly rising articulation function of Group III seems to be responsible for the interaction of Groups X Conditions reported in Table III.

Since the bisensory articulation function of Group

---

[1]Hallowell Davis, (ed.) Hearing and Deafness (New York: Murray Hill Books, Inc., 1947), p. 151.

III more closely simulated that of Groups I and II, it is suggested that the deviation of the auditory articulation function of Group III is responsible for the observed interaction of groups and conditions. Figure 3 may more clearly show the effects of the deviation of the auditory articulation function of Group III. Reference to that figure reveals that the parameter representing difference scores for Group III diverges most widely from those of Groups I and II at the sensation levels of 6 db and 15 db. It is precisely at these two sensation levels that the Groups X Conditions interaction attains statistical significance.

The Black Lists also produced auditory intelligibility scores in keeping with the findings reported in the literature. Black[1] reported an increase in intelligibility scores of approximately 9% per 4 db gain in intensity. This can be compared with the increase of approximately 8 - 9% per 3 db gain for Groups I and II of the present study. For Group III, the increase was approximately 6% per 3 db gain in intensity.

Group III again failed to improve in intelligibility to the same degree as Groups I and II as a result of increases in intensity. This strengthens the basic concept discussed previously, namely, that hearing losses of a sloping configuration tend to limit not only the maximum auditory recognition scores but, in addition, tend to affect the slope

---

[1] J. W. Black, op. cit.

or steepness of the articulation function.

In general, the auditory intelligibility scores earned with the Black Lists were higher than those earned with the PB Lists. The higher auditory scores obtained with the Black Lists appear as a result of two features of the lists. First, the vocabulary from which a response was drawn was quite limited. Sumby and Pollach[1] pointed out that the smaller the vocabulary from which a response is drawn, the higher the intelligibility score. Secondly, the Black Lists contain many two-syllable words. Mason's[2] study, designed to predict the intelligibility of speech, revealed that a word was more intelligible if it contained more than one syllable. Black[3] corroborated the work of Mason,[4] and reported that two-syllable words are more easily understood than are one-syllable words.

Certain features of the auditory performance among groups, however, were found to be similar for the two types of word lists. First, with each of the list types, the auditory articulation function of Groups I and II were sim-

---

[1]Sumby and Pollach, op. cit.

[2]H. M. Mason, Phonetic Characteristics of Words as Related to Their Intelligibility in Aircraft-type Noise, ORSD Report 4681. (Office of Technical Services Department, PBL 12160). 1945.

[3]J. W. Black, Accompaniments of Word Intelligibility," Journal of Speech and Hearing Disorders, XXII (1952) pp. 409-418.

[4]Mason, op. cit.

ilar. Secondly, the articulation function of Group III differed from that of Groups I and II with respect to its slope or steepness. The effect of the slower rise in the articulation function of Group III is evident in Table 5, where the differences among groups are statistically significant at the 15 db sensation level. The difference among groups may be attributed to the limited maximum auditory intelligibility scores earned by Group III. The bisensory recognition scores of all three groups tended to increase, although to a more limited degree at the higher sensation levels, as a function of intensity gain. It is apparent, then, that the failure of the auditory articulation curve of Group III to continue to rise, as did the curves of Groups I and II, accounts for the difference among groups. With the Black Lists, as in the case of the PB Lists, the auditory articulation function appears to be the basis of differentiation among groups.

Visual Recognition Scores

It may be recalled that the visual recognition scores of Group III differed to a statistically significant degree from those of Groups I and II when the Black Lists were presented, and that when the PB Lists were presented, a difference of statistical significance occurred between visual scores of Groups I and III. These findings indicate that with respect to speechreading ability, a sampling error was not avoided in the selection of subjects for Group III of the study.

The occurrence of such a sampling error was not completely unexpected, since all subjects in both hearing impaired groups were selected from the files of the University of Oklahoma Speech and Hearing Clinic. It is to be remembered that Groups II and III differed only in terms of the shape of the audiometric configuration. The configuration of Group II was relatively flat, while that of Group III was sloping. The flat configuration of Group II, together with the fact that the majority of the subjects in this group exhibited primarily conductive-type hearing losses, would suggest that individuals with similar configuration and type of hearing loss should make good use of a hearing aid. Individuals who make good use of a hearing aid do not usually seek the services of a speech and hearing clinic. In fact, a substantial number of members from the experimental Group II were, in reality, referred to the hearing clinic for routine audiometric tests because they were being considered as candidates for some form of middle-ear surgery.

With reference to Group III, however, the sloping configuration, combined with the fact that the majority of the group exhibited sensori-neural type hearing losses, suggests a somewhat limited ability in auditory discrimination. It is not altogether inconceivable that individuals who experience difficulty in the auditory discrimination of speech, and who in addition, may be poor speechreaders, represent the population of individuals who seek the services of speech and hearing clinics in an effort to improve their communication skills. Conversely, individuals possessing hearing losses similar to those of the subjects of Group III, and who are good speechreaders, may communicate more readily than the poorer speechreaders, and therefore, may not seek clinical help.

Bisensory Recognition Scores

The mean bisensory recognition scores were significantly better than the auditory scores. This was true for each group of subjects and at each sensation level tested when the PB Lists were used. The findings are in keeping with those reported in the studies of O'Neill[1] and of Sumby and Pollach.[2]

When the PB Lists were presented it was found that at the sensation levels of -6 db to 6 db, the increase in intelligibility resulting from the addition of visual cues had the same effect on Groups I and II as did the addition of approximately 6 db of intensity to the pure auditory presentations. For Group III, however, the addition of visual cues had the same effect on word recognition scores as did the addition of 9 db of intensity to the pure auditory presentations of -3 db to 3 db sensation levels. These

---

[1] O'Neill, op. cit.

[2] Sumby and Pollach, op. cit.

relatively consistent relationships imply that the use made of the bisensory cues of speech is more closely related to the use made of the auditory cues of speech than to that made of the pure visual cues of speech. In short, the degree of improvement at a given sensation level arising from a bisensory presentation was not constant or in keeping with the magnitude of the visual word recognition score, but rather, varied as a function of the intensity level of the auditory component of speech. Thus, it was impossible to predetermine bisensory performance by simply adding the visual score to the auditory score. However, a reasonable approximation of the bisensory performance at a given sensation level (within limits tested) would be the auditory score at a sensation level 6 db higher for Groups I and II or 9 db for Group III. It is quite apparent, then, that the bisensory performance is a result of an interaction between the visual and auditory cues of speech.

Differences among groups lend further support to the notion that the interaction between vision and audition is more closely related to auditory than to visual performance. Group III, it will be recalled, showed the greatest improvement with the bisensory presentation and differed from Group II in that the audiometric configuration was sloping rather than flat. If the interaction were not more closely related to auditory performance, we would not expect Group III to make a superior degree of improvement in intelligibility going from an auditory to a bisensory presentation, especially since those in Group II were superior speechreaders.

It may be impossible to state specifically the factors which allowed Group III to make significantly greater improvement than Groups I and II in the intelligibility of speech under a condition of bisensory stimulation as opposed to a purely auditory presentation of the speech stimuli. Nevertheless, there is reason to believe that a lesser degree of redundancy in the bisensory signal may be one of the factors working in favor of Group III. By referring to Figure 4, it can be seen that as the intensity of the auditory component of the bisensory signal is increased, the difference between the arithmetic sum of the auditory and visual scores and the bisensory score diminishes, indicating that each of the sensory signals (auditory and visual) are supplying, in part, identical information. It appears, then, that as the auditory discrimination increases, the visual cues of the bisensory signal become relatively less important in determining the maximum intelligibility of language presented at a given sensation level. Since the auditory discrimination of Group III is consistently poorer than that of Groups I and II, it follows that the visual signal is able to continue to supply a greater amount of additional information to the individuals in Group III.

The auditory and visual recognition scores obtained with the Black Lists were always higher than those obtained with the PB Lists. In the case of the bisensory presentations, however, the bisensory scores earned with the Black Lists were better than those of the PB Lists only at the lower intensity levels, but tended to equal scores of the PB Lists at the higher sensation levels. It will be recalled that the mean bisensory scores obtained with the Black Lists were never appreciably larger, and most often considerably less, than the sum of the auditory plus the visual recognition scores. Following O'Neill's[1] line of thinking, it is suggested that there is a greater degree of redundancy in the recognition of sensory signals which are available when the Black Lists are presented than occurs when the PB Lists are presented. Redundancy, as the term is used here, means that given sounds may be identified through their auditory pattern in isolation as well as by their visual movements. Thus, the visual cues and the auditory cues, each in isolation, provide for adequate recognition of the speech sounds, and therefore, a bisensory presentation of the same sound can do no more in terms of allowing for a correct identification than can an isolated visual or auditory presentation. The Black Lists represent a multiple-choice type of test, in which the subject must select his answer from only four possibilities. Since the vocabulary from which the choice

---

[1]O'Neill, _op. cit_.

is made is limited, the amount of auditory or visual information needed to make a correct identification is considerably reduced. It appears that the limited vocabulary of the Black Lists tends to foster redundancy, since relatively fewer visual or auditory cues are needed to make a correct response than when using a relatively unrestricted vocabulary. Figure 7 might be considered a graphic indication of the relative degree of redundancy for the Black Lists. Here the sum of the visual and auditory scores are compared with the bisensory scores, and the differences between them plotted. As can be seen, the arithmetic sum of the visual and auditory scores is most often greater (indicated by negative quantities) than the bisensory score. When the differences observed in Figure 7 (Black Lists) are compared to those of Figure 4 (PB Lists) it is noted that the bisensory scores for the PB Lists are most often greater than the arithmetic sum of the visual and auditory scores, implying, therefore, that there is considerably less redundancy operating in the recognition of sensory signals when presenting the PB Lists.

There is another aspect in which the bisensory recognition scores obtained with the Black Lists differ from those obtained with the PB Lists. It is in the steepness or slope of the bisensory articulation curve. The articulation function of the Black Lists is considerably flatter than that of the PB Lists. This may be due in part to the fact that the minimum bisensory scores are higher (in percentage of words correct) for the Black Lists than for the PB Lists. The higher minimum scores are felt to be partially due to the probability of selecting by chance the single correct response out of four possible responses. The higher word recognition scores for the limited vocabulary of four word (Black Lists) than for the comparatively unlimited vocabulary class (PB Lists) is in keeping with findings of Sumby and Pollach.[1]

For the hearing impaired population which seeks the help offered by lipreading instruction and auditory training, these findings have practical implications. It appears that in the rehabilitation of the acoustically handicapped through lipreading and auditory training, the two types of training should be combined into a unified instructional process. To teach lipreading in isolation and to establish auditory training as a separate entity seems unreasonable in light of the fact that most acoustically handicapped individuals retain some residual hearing, and therefore function with the aid of audio-visual cues of speech. Furthermore, since the purpose of training is to develop the maximum ability to receive verbal communication, advantage should be taken of the fortunate interaction which occurs between the auditory and visual cues of speech.

To make the assumption that training in the auditory

---

[1]Sumby and Pollach, op. cit.

discrimination of speech sounds apart from training in the visual recognition of speech sounds will enable the individual to achieve a level of performance equal to that attained by bisensory training seems hazardous. This is true because of the finding that the bisensory performance of the hearing impaired is not the simple arithmetic sum of the visual and auditory performances measured in isolation.

The present findings raise an important possibility for further research. It may be recalled that individuals in Group III exhibited a mean hearing loss for pure tones which dropped at a rate of approximately 12 db per octave interval in the frequency range 500 cps - 2000 cps. What is not known, however, is the magnitude of the drop in hearing acuity per octave interval that must exist before the performance between individuals with flat hearing losses and those with sloping losses would differ.

CHAPTER VI

SUMMARY

The present study explored the role of visual and auditory cues of speech and their interaction in the reception of verbal communication. Previous work in this area was either unrelated to the clinical population with impaired hearing, or when aimed in this direction, was poorly controlled. In either case, previous information failed to clarify the effects of such variables as hearing acuity, type of speech sample, and intensity of the auditory component upon the reception of speech.

The present study endeavored to clarify the effects of these variables by seeking information regarding the following specific questions:

1. How much does speech intelligibility improve as a result of combining the visual cues of speech with the auditory cues of speech?
2. How does the contribution of visual cues to the intelligibility of speech vary as a function of the type of speech material presented to the listener?
3. What effect does hearing acuity and configuration

of hearing loss have upon the use made of bisensory cues of speech?

4. How does the contribution of visual cues vary as a function of the intensity level of the auditory cues with which they are combined?

These questions were explored through an experimental design involving three groups of subjects. Group I consisted of individuals with normal hearing acuity, Group II contained subjects with moderate hearing losses with relatively flat audiometric configurations, and Group III was composed of listeners with moderate hearing losses with sloping configurations. Visual, auditory and bisensory word intelligibility scores were obtained for each listener. Under the visual presentation, only the visual cues of speech were available to the listener, while under the auditory presentation, only the auditory cues were available. Under the bisensory condition, both the visual and auditory cues of speech were available to the listener. The auditory and bisensory presentations were made at discrete sensation levels ranging from -6 db through 15 db in 3-db steps. Word recognition scores were obtained for two types of speech materials: phonetically balanced lists of words and multiple-choice intelligibility lists.

The data were analyzed in terms of group differences, differences in performance on various types of speech material, differences in performance resulting from variations in the intensity level of the auditory component of speech, and differences among methods of presenting the speech material. The analysis revealed the following information:

1. For the auditory presentations, the performance of the three groups was similar at the lower sensation levels tested. However, at the higher sensation levels, the performance of Group III (sloping loss) was appreciably poorer than that of Group I (normal hearers) and Group II (flat loss).
2. For the bisensory presentations, the three groups responded similarly at the lower sensation levels tested. At the higher sensation levels, the performance of Group III (sloping loss) was only slightly poorer than that of Groups I (normal hearers) and II (flat loss).
3. The bisensory word recognition scores earned with the phonetically balanced lists of words were always greater than the arithmetic sum of the visual score plus the auditory score.
4. The bisensory word recognition scores earned with the multiple choice intelligibility tests were most often less than the arithmetic sum of the visual score plus the auditory score.
5. When the pure visual presentations were made, Group III earned poorer word recognition scores than did Groups I and II.

6. When the performance on the bisensory presentations was compared to that of the auditory presentations, Group III made greater improvement than Groups I and II in the understanding of speech resulting from the addition of visual cues. This greater improvement cannot be attributed to superior lip-reading ability.
7. The amount of improvement in intelligibility when going from an auditory to a bisensory reception of speech, varies depending upon: (1) the type of speech material used, (2) the configuration of hearing acuity of the listener, and (3) the intensity level of the auditory component.

It is apparent, then, that there is an interaction between the auditory and visual cues of speech, since bisensory scores were not consistently equal to the sum of the visual and auditory scores. The degree of improvement in intelligibility owing to this interaction was greater for subjects with sloping losses than for normal hearers of those with flat losses. Moreover, the amount of information received from the visual and auditory cues in isolation will vary depending upon the type of speech material present, with greater redundance in the multiple-choice material.

These findings seem to warrant the use of a bisensory approach in the training designed to improve communication skills among the hearing impaired.

BIBLIOGRAPHY

American Standard Sound Measurement Package, American Standard Specifications for Speech Audiometers, (July 10, 1953). American Standards Association, Inc., New York, New York.

Black, J. W. "Accompaniments of Word Intelligibility," Journal of Speech and Hearing Disorders, XXII (1952), pp. 409-418.

Black, J. W. "Multiple-Choice Intelligibility Tests," Journal of Speech and Hearing Disorders, XXII (1957), pp. 213-235.

Davis, H., (ed.) Hearing and Deafness. New York: Murray Hill Books, 1947.

Edwards, Allen L. Experimental Design in Psychological Research. New York: Rinehart and Co., Inc., 1950.

Egan, J. P. "Articulation Testing Methods," Laryngoscope, LVIII (1948), pp. 955-991.

Ewing, I. R. Lip Reading and Hearing Aids, Manchester, England: Manchester University Press, 1944.

Hutton, C. "Combining Auditory and Visual Stimuli in Aural Rehabilitation," Volta Review, LXI (1959), pp. 316-319.

Jerger, J. F., Carhart, R. T., Tillman, T. W., and Peterson, J. L., "Some Relations Between Normal Hearing for Pure Tones and for Speech," Journal of Speech and Hearing Research, (II) (1959), pp. 126-140.

Mason, H. M. Phonetic Characteristics of Words as Related to Their Intelligibility in Aircraft-type Noise. ORSD Report 4681. Office of Technical Services Department, PBL 12160.

Neely, K. K. "Effects of Visual Factors on Intelligibility of Speech," Journal of the Acoustical Society of America, XXVIII (1956), pp. 1275-1277.

O'Neill, J. J. "Contributions of the Visual Components of Oral Symbols to Speech Comprehension," Journal of Speech and Hearing Disorders, XIX (1954), pp. 429-439.

Sumby, W. H., and Pollach, I. J. "Visual Contribution to Speech Intelligibility," Journal of the Acoustical Society of America, XXVI (1954), pp. 212-215.