FINAL TECHNICAL REPORT - SDC 411-1-17

INTELLIGIBILITY IN VOICE COMMUNICATION

(Psychological Studies of Training Techniques)

Kenyon College
Gambier, Ohio
22 September 1949

SDC Human Engineering Project 20-K-2
Contract N7onr-411, T.O. I
Project Designation NR-782-004

Investigations conducted at:
Acoustic Laboratory
School of Aviation Medicine and Research
N.A.S., Pensacola, Florida

John W. Black
Project Director

Ashton Graybiel
Captain (MC), USN
Technical Officer

Report prepared by:
John W. Black

FOR THE SPECIAL DEVICES CENTER:

Reviewed for Human Engineering Branch:

J. Gaberman, Project Engineer
Code 913

Submitted:

C. S. Rhoads, Technical Director

H. A. Voss
Head, Code 913

Approved:

Page Knight, Commander, USN
(Acting) Commanding Officer & Director

Table of Contents*

*The numbers in parentheses refer to the technical reports of the project. The titles appear at the end of the report. Sections that include topics that are discussed for the first time in the work of this contract are underlined.

<u>Introduction</u>. Voice communication, the subject of the researches of this contract, involves a series of assumptions. This is particularly true with regard to military applications of voice communication. First, there is a general opinion that face-to-face talking is efficient, that meaning is adequately phrased and spoken by one person and picked up in identical form by as many listeners as happen to be in his view. Evidence to the contrary in the living room, automobile, or office rarely changes the prevelant belief that 'talking is natural', 'talking is easy', 'talking is efficient.' Second, and as a part of the communication chain, the view prevails that everyone listens alike. Tests of individual differences among listeners establish that this ability is distributed in about the same manner as other physical and mental traits. Beyond the fact that assumptions abound in the phrase <u>voice communication</u>, there is the strong pragmatic view that 'talking has worked', it must be good. This could be questioned on other philosophical grounds.

The matter of importance is that voice communication, imperfect in normal circumstances, is relied upon in military operations where both speed and accuracy are needed. An airplane traveling 600 m.p.h. goes 170-200 feet during the time that it takes a person on the ground to say <u>cop</u> conversationally. But conversational standards do not apply where the listener-talker is surrounded by noise levels that tax the well being of the ear. The intensity of the voice signal must be increased and concomitantly the duration of the syllable is lengthened.

The objective of the present researches has been improvement of voice communication. As a basis for this, studies have been directed to describing speaking in a variety of circumstances, particularly in different acoustic environments. Variation in normal speech is found to accompany different (1) types of messages, (2) sounds, and (3) stimulus voices. Whether these are cultural effects or not they affect voice intelligibility, one segment of the term voice communication.

The purpose of the present report is to synthesize the results of the experiments of the contract. The details of experimental procedures are omitted, and the conclusions are summarized briefly.

1. Variables in normal speech.

A. The vowel. The principle intensity and pitch of the word or syllable are contributed by the vowel. Differences in the relative intensities of the vowels have been reported; also average differences in fundamental pitch. The study that is summarized here was an attempt to measure intensity, fundamental frequency, and durational aspects of the same samples of speech and to determine inter-relationships among these variables.

Forty-two subjects (Ss) read 11 monosyllables that contained as many different vowels. The readings were recorded both phonographically and with a power level meter (Sound Apparatus). Measures of the mean frequency, duration, and intensity components were made and analyses of variance were performed on the readings of the 16 Ss who were successful in making all of the vowels. (The responses of these 16 were not significantly different from those of the entire group of 42 readers in any measure.) F was significant in all of these analyses. The arrays of the mean values appear in Table I. Of the 55 possible comparisons in each instance the following numbers of pairs were highly significant

(1%): frequency, 25; duration, 32; and intensity, 9. Both frequency and duration tended to be related to the 'openness' of the vowel: the more open the vowel the lower the frequency of the fundamental and the greater the duration of the sound.

B. Types of messages. Two experiments were conducted in which comparisons were made of the intensity and durational characteristics of different kinds of messages. The studies complemented each other in that different messages were used in the various conditions in one experiment; in the other, identical phraseology was used from one condition to another. In the first experiment 48 Ss spoke six short messages: (1) reading, (2) repeating, (3) continuing the first part of familiar statement, (4) locating a familiar site, (5) reading responsively, and (6) describing a picture. Most of the phrases contained five syllables. The order of presenting experimental conditions was counterbalanced. The mean of the four peak intensity values of a phrase was used as the intensity measure for the phrase as spoken by one S. The mean of the six phrases/condition was treated an an S's response to a condition and as a basic measure in an analysis of variance. The comparable value for the analysis of duration was the mean duration of six phrases of a condition as spoken by one S.

The results of the analyses of variance appear in Table II. Both intensity and duration differed from one type of speaking to another. The range of syllabic duration approached a 2:1 ratio. Reading and repeating were the slowest types of speaking (three syllables/sec.); and locating sites and reading responsively, the fastest (five syllables/sec.)

In the comparable study 72 Ss spoke five short phrases as messages read (1) directly and (2) responsively, and (3) as impromptu responses to questions. The Ss participated in three separate groups on as many days, and the order of conditions was rotated among the groups.

Difficulties with equipment made the data difficult to interpret. These were partially overcome by measuring the duration of the phrase at a constant level with respect to the peak intensity of the phrase (-25 db). The F-ratio between type of message and remainder (message type subjects) was highly significant. The means enumerated in Table III show reading as a relatively slow type of performance as in the experiment summarized above. Intensity measures for the first group of Ss in this study were similar to those cited in the earlier experiment. Lack of significance with the other two groups was possibly attributable to mechanical failure or to questionable experimental design. With regard to the latter alternative the possibility arises that differences in intensity that accompany various messagettypes are sufficiently subtle that they might not appear in 'repeated' phrases. Possibly a phrase that is said a second time is predominantly a 'learned' phrase, not an example of 'description' or impromptu speech. The fact, however, that results with duration in this experiment—and with one group of Ss, intensity—were similar to those of the earlier one would indicate that the mechanical difficulties were probably the main reason the results only partially corroborated the ones in the companion study.

C. Speaker variability. The two preceding topics show that voice messages differ. Some of the variability might be attributed to relatively fixed relationships between the physiology and the acoustics of talking. For example, 'the more open the vowel, the greater the duration' might seem to be an inevitable relation. The lower jaw makes a greater excursion in forming an open than a closed vowel. This in turn, takes increased time--assuming that in normal speech there is a relatively fixed rate of jaw movement for a single speaker. Plausible as this is, the series of studies has shown significant individual differences in every experiment. Approximately 100 analyses of variance have been performed. Data have been contributed by over 2000 Ss. The normal experimental unit has been a multiple of 12 Ss. More frequently than not the experimental design has included repeated measures from the same Ss. In every analysis involving voice the variance attributable to individuals has been statistically significant.

2. Testing voice communication.

A. Intelligibility tests. The multiple-choice intelligibility tests that were developed by the NDRC Voice Communication Laboratory are used by the Services for testing intelligibility in voice communication courses. Also the tests are used some in research. Practice varies with respect to the tempo of reading the speaker lists of the tests. Each of eight items has three parts and the whole is scored on a basis of 24 items/speaker. In all there are 24 speaker tests. The listener's answer form permits a choice from four words for each one that is spoken. Responses, i.e., crossing out the words that are heard, are made rapidly. However, it has not been clear whether a rapid reader penalizes his score by getting ahead of his listeners, quite apart from his intelligibility.

Twenty-four Ss read as many speaker lists, one list/speaker. Reading was slow. With the aid of a sweep-hand time S paused 4 sec. between items (phrases) and 2.5 sec. between words within an item. The reading of the tests was recorded with a high-fidelity microphone and a magnetic tape recorder. S read in quiet. The tape was edited 17 times in order that inter-phrase pauses of 2, 3, and 4 sec. occurred respectively with 2.5, 2, 1.5, 1.0, .5, and 0 sec. inter-word pauses. Each edition of the tape was copied to a disc recording that was played back to a panel of listeners (9-12) in noise (105 db). An analysis of variance was made of the error scores of the listeners. This analysis and the means of the 18 conditions appear in Table IV.

The large variance attributable to phrase-word interaction and the lack of systematic order within the columns of means suggest experimental error. In spite of this the word interval variance is significantly greater than the largest interaction value. However, the main problem under test was not definitely answered. Possibly the noise level in the testing room varied from session to session, although it was adjusted at the outset of each period for 105 db (General Radio). Possibly the copying of the magnetic tape to discs was more efficient in some cases than in others. Within the limits of the experiment no advantage was shown for pausing longer than 2 sec. between phrases in reading the test. The effect of varying rates within a phrase was not clear, although it appeared that speed of reading a three-word phrase was not critical in intelligibility testing. (A study is planned in which the same voice will read all of the tests and the variable will be limited to word interval.)

A second aspect of intelligibility testing that has been investigated is the effect of interposing a recording-reproducing system between the speaker and the listener. Six groups of Ss (12 each) read the multiple-choice intelligibility tests. Each group provided a panel of 11 listeners as the 12 members read speaker lists in rotation. Two voice recordings were made simultaneously with the 'direct' intelligibility testing, one on discs and the other, magnetic tape.

All testing was done over an airplane 'intercom' mock-up with listeners and speakers in 105 db of airplane-type noise. The recordings were played back to different listening panels and the scores assigned the same speakers by the three methods were compared. There were 18 listening panels. In an analysis of variance (Table V) significant difference was attributable to conditions: (a) direct *vs.* (b) disc recording *vs.* (c) tape recording. The mean error scores for the respective conditions were 7.42, 8.53, and 8.55 (of 24). Both recording conditions differed significantly from direct scoring (t, 1%, any difference between means of .77). In practice, Ss scores are determined in percent right. In the present instance values would be respectively: 69.1, 64.5, and 64.4. Product-moment correlations between the arrays of scores of individual speakers graded by the three different methods were as follows: r, a-b, .83; a-c, .81; b-c, .88. These correlations are as high as are claimed for the tests themselves. The conclusion is that relative intelligibility scores can be assigned as well with high-fidelity recording equipment in the speaker-listener circuit as with listeners scoring a speaker directly. An individual score, however, is penalized approximately 5% by the recording process.

B. Materials for measuring speech. In the investigations of voice under specific conditions and for purposes of comparisons control of the stimulus materials beyond ordinary precautions was indicated. Exploratory studies showed that messages that were equated in numbers of syllables and appeared equivalent might vary inherently in both duration and intensity values. The data related to vowels, for examples showed that this variability extended to the frequency of the fundamental in vowel sounds.

An early series of experiments employed as standard stimulus materials five lists of 12 five-syllable phrases each. These were selected from recorded R/T procedures. The lists were later edited to eight phrases, equated in intensity and duration characteristics, and used as materials for reading and repeating.

As a basis for more general measurements four hundred ninety-two phrases were selected from <u>Navy Flight Patter</u> and grouped into sub-lists of 12 each. Critera in the selection of a phrase were: five syllables, 3-5 words, and maximum of two syllables for any word. Eighteen Ss read all the phrases. The mean duration of the phrases ranged from .88 to 1.54 sec. and in standard deviation from .03 to .40. Significant differences in mean duration accompanied (1) sub-lists and (2) number of words/phrase. The same phrases were subsequently studied for natural intensity. Thus it became possible to draw up lists of test phrases equated in intensity and duration. These were used for securing repeated measurements from the same Ss without their speaking identical contents in different performances.

In determining these measures readings were taken from different meters. The correlation between the intensity measurements provided by the General Radio sound level meter (slow) and the Sound Apparatus power level recorder exceeded .9. Also there was high correlation (r = .96) between the intensity of phrases as indicated by (1) peak intensity and (2) the mean of four peaks.

3. Controlling speech through spoken stimuli. Various approaches have been made toward the improvement of the intelligibility of voice communication: (1) alteration of communication equipment, (2) protecting the listener's ear from masking voice, (3) standardization of messages, and (4) training personnel to use their voices to better advantage. Another approach has been explored in a series of experiments in which the effect of heard voice upon responding voice has been investigated.

A. Intensity. The best documented concomitant of intelligibility over communication systems in noise is strong vocal intensity. A loud voice, with proper use of equipment, contributes a favorable signal-to-noise ratio, a requisite for satisfactory listening. That adequate intensity for reasonably clear transmission can be acquired through brief training has been shown. The results, however, are averages for groups of Ss and are obtained during routine periods of intelligibility testing. They are not assuredly present with each intervening or succeeding voice transmission by each S Probably the intensity of the flight messages fluctuates when the talker is somewhat under the control of his listener, the two voices alternating in two-way communication. A resumé of the relevant findings: 25 Ss repeated 60 five-syllable procedural messages that were heard at five levels of intensi . -12 messages/level. The levels represented a range of 85 db at the input of the headset and from barely intelligible to very loud at the ear. A similar number of Ss spoke obvious one-word replies to 60 questions that were heard at the same five levels of intensity. The important difference between the two circumstances was that S invented the answers in one instance and not the other. In both cases, except when very weak intensities were under comparison, the means of the peak responses were higher as the intensity of the stimuli was increased. In a third part of the study 16 Ss were requested to maintain a single level of talking irrespective of the intensity of the stimulus. The tendency of reply in keeping with the intensity of the received messages persisted. The data of these experiments are summarized in Table VI.

In each of the three studies the responses to the weakest stimuli were numerically more intense than the response to level 2. The possibility arose that a significant effect was being obscured by the gross increment in the intensity of the stimuli (20 db) between these levels. An experiment was planned to test this possibility. One voice recorded five lists of eight words each. The first and last four words of each list were antonyms of each other, for example, full, narrow, sister, many, empty, wide, brother, few. S heard the list two times at five levels of intensity. On one occasion he repeated the words, and on the other he said the opposites of the words. In both instances he said the same words although in response to different sets of directions. The order of conditions was counter balanced and each list was used with each condition an equal number of times.

The five levels of intensity were six db apart and the lowest one was 5 db above S's word-discrimination threshold. This threshold was found by asking S to repeat successive 12-word lists heard at a succession of low levels. The words and lists were equated for intelligibility. The level at which his responses were 50% correct was called his 'word-discrimination threshold'. The results of the experiment are summarized in Table VII. The responses to level 4 were significantly (1%) less intense than the responses to all weaker stimulus levels. It is reasonable to suppose that some of these levels of stimuli, including level 4 fell between levels 1 and 3 in the earlier study.

There would seem to be two general ways of increasing the intensity of the responding voice: (1) give the ear of the listener-speaker very faint signals; or (2) give the ear very intense signals. Since the increments in the intensity of response were small in the first instance it has little practical application.

There was no difference in intensity corresponding with the repetitions-opposites comparison.

The pattern of the results of the experiments described above recurred several times in other experiments in which the effect of the intensity of the stimulus upon the intensity of the response was not the main experimental variable. In one instance four types of responses were required to stimuli of different levels of intensity. The stimuli were words--as described in the preceding paragraphs--and digits and letters. The spoken responses to the words were: (1) repetitions and (2) antonyms. The responses to digits (and letters) were:

(1) repetitions and (2) the solutions to simple problems--addition and subtraction. In all four instances the mean intensity of the responses varied with the intensity of the stimuli in the manner of the preceding discussion.

The results cited were derived from experiments in which there was a time lag between hearing the stimulus and saying the response. Other studies treated talking that occurred simultaneously with the experimental auditory stimuli, i.e., sidetone. In one experiment the equipment was an interphone mock-up with the amplification of S's sidetone varied systematically. Sixteen Ss read standard intelligibility tests, one with each of four levels of sidetone. Two were above and two below the normal level provided by the interphone amplifier. Speakers and listeners were surrounded by high-level noise (110-114 db). The intensity of the spoken messages and the intelligibility of the Ss were measured. As the amplification of the sidetone was increased the readers attenuated their levels of reading. Concomitantly their intelligibility was reduced. The results are summarized in Table VIII.

The disimilarity between this experiment and the foregoing ones is obvious, both in plan and results. In this case the more intense sidetone produced attenuated, not more intense speech. The two sets of studies were alike in that in both the intensity of the heard stimuli affected the intensity of the responses; they were dissimilar in the direction of the change. When the listener could alter (reduce) the intensity of the strong signal (sidetone) that he heard, he did so. When he was not speaking and could not change the signal level of the person who was speaking, he became intense with the stimulus that he heard.

In none of the studies of the effect of intensity (heard) upon intensity (spoken) could a mathematical function be determined that obtained for other studies. The quantification of this relationship is doubtless possible. It requires recording-reproducing equipment with greater dynamic ranges than was available; also an absence of noise in the system at all levels under comparison.

The question arises whether the listener-speaker responded to the intensity or the loudness of the stimulus. Only the former was controlled in the studies cited. An exploratory study was designed to find whether the general pattern, 'The more intense the stimulus, the more intense the response' could be extended to, 'the louder the stimulus....' Pure tones of 10 frequencies, 98-247 c.p.s., stimulated an S's ear for 10 sec. The intensity of the tones was constant. This meant, in terms of the Fletcher-Munson equal-loudness curves, that the loudness of the stimulus tones was dissimilar. After the 10 sec. stimulation S (24 Ss) read a nonsense syllable containing the vowel (ʌ). The intensity of the spoken syllable was measured. S read a total of 30 syllables, three responses to each of the 10 frequencies. An analysis of the results appears in Table IX. There were significant increments in the intensity of the reading accompanying the loudness levels of the stimuli. Possibly the precise function that would describe the effect of heard intensity upon spoken intensity would be in terms of loudness instead of intensity.

B. Rate. In a set of experiments similar to the ones conducted with intensity the effects of different rates of heard speech upon the rates of responses was investigated. Five-syllable phrases were recorded at different rates of speaking, varying in total duration/phrase, 1-4 sec. Groups of Ss heard and repeated the phrases under instruction to talk naturally. In one experimental plan the Ss heard five lists of phrases with all of the phrases of a single list spoken at the same rate--five rates for five lists; in another, the rates were varied from phrase to phrase. The duration of the responses was measured. In a subsequent experiment the same technique was followed except that the natural readings of five speakers with differing rates were used as stimuli. The results of the investigations are summarized in Table X. These three studies established a pattern, 'The faster the rate of the stimulus, the faster the rate of the response.' This pattern recurred in all studies involving time differences among stimuli.

One extension of principle was studied with respect to pauses in the stimulus phrases. Phrases were recorded (a) without pauses, (b) with natural pauses, and (c) with unnatural or illogical pauses. The mean responses were significantly different in duration: (1) the inclusion of any pause in the stimulus lengthened the responses; (2) the logical pauses increased the duration of the responses more than did the illogical ones.

No study yielded satisfactorily a quantitative statement of the relationship between vocal intensity and duration. Several studies gave indications that the two were positively related, and that as one was shaped by either S's physical or verbal environment there was an effect upon the other. An experiment was designed to try to determine the relationship. Questions of experimental error made the results somewhat less than definitive, although indicative of a close relationship between intensity level and rate. Eighteen Ss heard and repeated equated phrases at each of three rates and, in turn, from three transmission systems: voice (direct), earphones, and loudspeaker. The messages were heard at a low level. A second group of Ss repeated the procedure while hearing the items at a conversational level. And a third group heard the messages loudly.

The effect of the rate of stimulus on rate of utterance was clear in all comparisons. The effect of the intensity of the stimulus upon the intensity of the response was also clear, and in keeping with the results of earlier studies. The differential effects of the transmission systems were more difficult to assess. Table XI lists the mean intensity and duration values for the various conditions.

C. Pitch. Two studies were conducted to find whether the pitch (fundamental frequency) of the responding voice was affected by the corresponding attribute in the stimulus voice. In one, 60 five-syllable phrases were recorded with half of the inflections of the final syllable up and half down (random order). The same phrases were re-recorded with inflections opposite those of the first recording. Groups of Ss (24 each) heard and repeated the phrases. The repetitions, in turn, were recorded and the recordings played back to groups of judges. The judges indicated phrase by phrase whether the messages ended with an upward or downward inflection. A majority of judgments/phrase/speaker was taken as an indication that the phrase was spoken <u>up</u> or <u>down</u>. The results were evaluated by a chi-square technique. More than one third (23/60) of the phrases 'followed the stimulus' at the 1% level of confidence; and more than one half (35/60) at the 5% level. Definitely the pattern was for the talkers to 'say back' the messages with the inflection that accompanied the stimulus.

The second study is inconclusive at present. An investigation, discussed above, related to the intensity with which Ss responded to tones that represented different loudness levels (constant intensity). Frequency measurements were made of the oral responses to the different frequencies. There were significant differences among the means of the responses. Thus it might be interpreted that, 'The higher the stimulus tone (c.p.s.) the higher the response.' However, as noted above there were differences in vocal intensity accompanying the responses to the different tones. There is a positive relationship between vocal intensity and frequency, the more intense voice having the higher frequency (for the same individual). Therefore, the increments in frequency might have been (1) incidental accompaniments of the increments in intensity, or (2) <u>vice versa</u>, or (3) unrelated phonomena corresponding with c.p.s. and loudness levels of the stimuli independently. Further work is in process on this problem.

D. Articulation. Another demonstrated determinant of voice intelligibility is precision of articulation. Physically this is probably a function of amount and duration of breath pressure at the places of articulation. In an experimental situation Ss were found to speak in a manner to copy the degree of articulation that they heard immediately before talking. Briefly, five voices recorded 12 procedural messages each. Varying degrees of articulation were deliberately used by each reader with his over-all intensity and duration held constant. Twenty-four Ss heard and repeated the messages under the specific instruction to <u>talk naturally</u>. These responses were recorded phonographically. The repetitions of one well-articulated phrase and one poorly articulated phrase (as spoken by each stimulus speaker) were re-recorded and played back in pairs to panels of judges. Fifty-three judges selected the 'better articulated' message in each of the 120 pairs of repetitions. In more than half of the pairs of responses the judgments showed that the repetitions of the well-articulated phrases were uttered more precisely than the repetitions of the poorly articulated messages--this at the 10% level of confidence. Judgments of more than one-third of the pairs deviated from a chance outcome at the 1% level of confidence and in a manner to associate the precision of articulation of the responses with that of the stimulus. These results are summarized in Table XII.

E. Physical environment. Rooms. A second general method of influencing the manner in which a person speaks is to alter the physical conditions surrounding him. An experiment was conducted in which 184 Ss participated. Reading occurred in eight rooms. The rooms represented two sizes, shapes, and reverbration times. Twenty-three Ss read in each room. The same phrases were used throughout. Measurements were made of the intensity and duration characteristics of the reading. The results are summarized in Table XIII.

Both vocal rate and intensity in the saying of a series of short phrases were affected by the room conditions under which the reading occurred. Phrases were read slower in large rooms than in small ones, and among large rooms, the rate was slower in live than dead rooms. During reading a series of phrases the mean rate became faster, more sc in large than in small rooms. This result was computed by comparing the rate of the first and last three phrases as read by each S.

Intensity of reading was greater in dead than in live rooms, particularly in the larger ones. This interpretation depends upon the accuracy of a directional microphone to react almost exclusively to a direct wave front emanating from a talker and the supporting observation that during the reading of the phrases S's vocal intensity increased in dead rooms and either decreased or remained constant in live ones. More intensity activated the microphones in small than in large rooms. It is not clear that this resulted from corresponding differences in vocal output; it may have resulted from reflected wave fronts.

The data indicated that the reader monitored his sidetone as he read and adjusted his vocal behavior during reading. This response to 'feed back' was consistent with maintaining a 'normal' experience at the ear.

F. Cubicle. As an extension of the study of rooms an investigation was made to find whether the difference between large and small rooms carried further to small temporary enclosures. A group of Ss read groups of equated phrases in noise and quiet and in each instance with and without a 2 x 2 x 2 ft. cubicle about the head. The room and cubicle were sound treated. S spoke into a carbon lip microphone (M-6/UR). An analysis of variance failed to show any difference in intensity in the cubicle-no cubicle comparison.

G. Illumination of a room. An experiment was conducted in which 30 Ss sat in three degrees of illumination (bright, half-light, and dark) while repeating recorded word lists. No difference in word intensity occurred in this comparison.

H. Noise. The analysis of the Cubicle (F) experiment did show significance with regard to levels of reading in noise-no noise. Table XIV summarizes the results of the study. Exposed to noise (105 db) S spoke nine db more intensely than when in quiet.

I. Microphone placement. A study of the effect of microphone position on vocal intensity in quiet was inconclusive, due to lack of success in repeating the experiment. Sixteen Ss took part in the study. Each sat with his head in a constant position, a fixed distance from a hidden microphone. S read identical passages while an experimenter (E) shifted a dummy microphone 1, 2, 3, and 4 feet from S. Intensity measures were secured from the hidden microphone. The results are summarized in Table XV. When the experiment was repeated with a different order of conditions no significant differences were observed. This, however, was in combination with an error in the calibration of the equipment. The level of response of the system was attenuated from the level used when the results above were obtained. On another occasion when the output of the microphone amplifier was reduced a similar—although less important—discrepancy in results appeared (see message types, above). It is possible that at low levels of response the circuit did not discriminate among small differences in input intensity. Another result that was inconsistent in the two trials of this experiment was the difference in intensity when S was directing his speech to a live listener (E) who was alternately four and eight feet from S. In the first trial S was significantly more intense when speaking over eight feet. In the second trial there was no difference. In both trials of the experiment S was significantly more intense when speaking to a microphone than he was when speaking to a person--both four feet from him.

J. Audience. Two groups of Ss of different military status, officers and officer candidates, read lists of phrases (1) alone, (2) to their fellows, and (3) to members of the 'other group'. The intensity and the duration of the phrases were measured. There was no significant difference in either the intensity or duration of the officers' reading under the three conditions. The candidates, however, varied in intensity, talking louder in the presence of an audience than when alone (Table XVI).

K. Training. Earlier experimentation has demonstrated the beneficial effects of training for voice intelligibility. Training syllabi stress the importance of Ss' speaking loudly into the microphone. This instruction--to talk loudly--is variously phrased. In an experimental situation typical wordings were used with Ss during a series of intelligibility tests. S, upon reporting to a central talking station, was handed a card on which was typed one of six messages (in addition to the lines "This is an intelligibility test. Follow the instructions on your speaker's card. Hold the microphone lightly touching your lips."

(1) IF YOU ARE TO BE HEARD, YOU MUST SPEAK JUST SHORT OF SHOUTING

(2) IF YOU ARE TO BE HEARD, YOU MUST SPEAK SO THAT YOU FEEL THE STRAIN OF SHOUTING.

(3) IF YOU ARE TO BE HEARD YOU MUST SPEAK IN A MANNER TO PRODUCE A GOOD CLEAR SIDETONE (IN YOUR OWN HEADSET).

(4) [NONE]

(5) (None, but in this condition, immediately after S spoke the first item of the test, E said, "Say again; speak louder.")

(6) IF YOU ARE TO BE HEARD, YOU MUST SPEAK LOUDLY.

Speaker test lists were counter balanced among the six conditions. Twenty-four Ss received each instruction.

An analysis of variance was made of the scores in each condition and of the intensity of the speaker. The results are summarized in Table XVII. Three of the directions were significantly more effective than none in securing increased intensity of reading, Conditions 1, 2, and 6. The fact that condition 3 was not significantly different from none in intensity and produced the highest intelligibility score is interesting. There is the possibility that the direction was interpreted more in terms of articulation--another accompaniment of intelligibility--than of intensity. The results indicate that for securing greater intensity from Ss, written directions using the words shouting or speak loudly are beneficial. Possibly good, clear sidetone contributes to precision of articulation.

4. Physiology.

A. Loud talking. A study is under way to determine the physiological effects of talking. As a point of reference 'loud reading' was used as talking performance. Twenty Ss read as loudly as possible for 30 min. Continuous oximeter readings were made during the reading and for 30 min. following the reading. Also alveolar air samples were collected each 10 min. Thus far the results of speaking seem to be those that attend hyperventilation.

B. Coryzal speech. An effort was made to quantify the effects of a cold upon speech. Ss read intelligibility tests 'with a cold', 'with a treated cold', and 'after recovery'. (1) Intelligibility was not systematically changed as a result of a cold. (2) From recordings judges were able to identify voices with a cold. (3) No single group of speech sounds seemed to distinguish a 'cold' voice.

C. Temperature. An attempt was made to determine the effect upon body temperature of sustained exposure to high-level noise. This study related to the assumptions that body temperature is indicative of sleep *vs*. wakefulness, and the question posed was whether noise plus the task of hearing and writing words contributed to sleep or wakefulness. The study was incidental to other researches that were in progress, specifically to the rating of large numbers of words in intelligibility.

Approximately 400 Ss participated in the study in the following manners and at the indicated times.

	Experimental	Control
	Exposed to noise while hearing and writing word tests	Engaged in clerical work in quiet
0700 - 0930	N,142	N,55
1200 - 1430	N,147	N,40

The mean temperatures of the four groups at the outset and end of the periods follow:

0700 - 0930

Experimental	Control
98.3 - 98.2	98.2 - 97.8
t, 2.56 (1%)	t, 6.25 (1%)

1200 - 1430

Experimental	Control
98.5 - 98.3	98.7 - 98.0
t, 4.07 (1%)	t, 8.78 (1%)

The mean temperature of all groups fell significantly during the periods under consideration. However the temperatures of the control groups fell significantly more than the experimental groups (A.M., t, 3.75; P.M., t, 5.88). Thus if these measures were taken during normal times for temperatures to fall for these Ss the noise condition may be thought of as retarding the trend. The element of least control in the experiment was probably the amount of work that the Ss were doing. Writing word lists is a compelling activity that is not identical with clerical work.

5. Listening.

A. Monaural listening. Little work has been done under the contract with listening apart from speaking. In one experiment monaural and binaural listening were compared. There were indications that monaural reception was superior. When this was tested further with a small number of Ss participating in a Latin square experimental design no difference was isolated. As one part of this study some Ss were exposed to high-level noise for 1-1/2 hours. This did not affect their ability to hear recorded intelligibility tests in noise.

B. Posture. In a second study the effect of head posture upcn listening ability was tested. No relationship was found. Ss heard equally well in any of four postures: head erect, head forward, head left, and head right.

Table I

Mean values of frequency, duration, and intensity measurements of 11 vowels. N, Ss, 16.*

	Frequency (c.p.s.)	Duration (sec.) Level recorder	Duration (sec.) Magnetic tape	Relative intensity (db)
(i)	145.7	.252	.159	0.00
(I)	141.7	.251	.135	2.86
(e)	136.5	.286	.191	1.77
(ε)	137.6	.265	.153	3.12
(æ)	132.5	.331	.208	3.44
(ɑ)	132.7	.317	.192	3.69
(ɔ)	134.8	.318	.209	3.22
(ʌ)	140.5	.263	.154	2.21
(o)	137.0	.269	.197	3.71
(ʊ)	148.8	.248	.153	2.52
(u)	153.0	.292	.200	2.56

*Any difference between means significant (t) if it exceeds:

	t,1%	t,5%
Frequency	7.3	5.5
Duration		
Level recorder	.032	.025
Magnetic tape	.028	.021
Intensity	2.17	1.65

Table II

A. Summary of analyses of variance oi intensity and duration measures of six types of speaking. Rate: basic measures, mean syllabic duration (sec.) of approximately 30 syllables/condition/S. Intensity: basic measures, mean intensity (db) of 24 peaks, four in each of six phrases/condition/S. N, Ss, 48.

Source of variation	d.f.	Mean square rate	Intensity
Conditions	5	592.65**	362.40**
Subjects	47	116.94	306.30
Remainder	235	8.14	2.54

** Significant (F) at the 1% level of confidence.

B. Relative means of values analyzed in A.

	Duration* (sec./syllable)	Intensity** (db)
1. Reading	.33	4.4
2. Repeating	.37	3.7
3. Continuing statements	.26	0.0
4. Locating sites	.20	1.0
5. Reading responsively	.20	2.2
6. Describing	.27	3.8

* Values derived from graphic level recorder readings. Any difference between means of .021 significant (t)at the 1% level of confidence; .016, 5%.

** Intensity values relative to the least intense condition, i.e., condition 3 = 29.2 db = 0; this is comparable to 68.5 db (General Radio). Any difference between means of .830 significant (t) at the 1% level of confidence; .631, 5%.

Table III

Mean durational values of three types of messages, measured 25 db below the peak intensity of the phrase. N, Ss, 72.

Type of speaking	Duration (sec./phrase)*
Reading	3.49
Responsive reading	3.23
Impromptu speeches	3.06

* Difference between means of .25 significant at the 1% (t) level of confidence; .19, (5%).

Table IV

A. Summary of analysis of variance of error listening scores of successive panels of listeners who heard the same intelligibility tests with specific inter-item and inter-phrase pauses.

Source of variation	d.f.	Mean square
Phrase interval (P)	2	8.7
Word interval (W)	5	277.9*
Tests (also speakers) (T)	23	59.5
P x W	10	51.5
P x T	49	1.1
W x T	115	3.4
P x T x W	230	1.2

* Significant (F) at the 5% level of confidence.

B. Mean error scores analyzed in A.

Word interval (sec.)	Phrase interval (sec.) 4	3	2	Over-all
2.5	11.9	9.9	14.1	10.2
2.0	8.1	7.2	6.9	6.3
1.5	5.9	7.3	4.1	4.9
1.0	7.1	9.6	10.2	7.6
.5	9.6	8.0	8.1	7.3
.0	13.0	11.7	8.8	9.5
Over-all	7.9	7.6	7.4	

Table V

Summary of an analysis of variance of the intelligibility scores assigned to 72 speakers through (a) immediate listeners, (b) listeners hearing disc recordings of the tests, (c) listeners hearing tape recordings of the tests.

Sources of variation	d.f.	Mean square
Speakers	71	16.64
Transmission systems	2	30.43**
Remainder	142	1.03

** Significant (F) at the 1% level of confidence.

Table VI

A. Summary of analyses of variance of the intensity of oral responses to stimuli heard at five levels of intensity. (1) Repetitions: single words, 25 Ss. (2) Answering questions: single words, 25 Ss. (3) Repetitions: single words at constant level, 16 Ss.

		Mean square		
Source of variation	d.f.	1	2	3
Intensity	4	502.75**	440.21**	62.72**
Subjects	24(15)	75.75	101.44	166.99
Remainder	96(60)	4.78	5.08	4.59

** Significant (F) at the 1% level of confidence.

B. Means (and SD's) of the values analyzed in A.

	Mean response (db)		
Stimulus	1	2	3
1. Minimal for understanding single words	74.02(4.67)	71.26(4.64)	71.25(6.11)
2. Condition 1 plus 20 db	73.14(3.73)*	70.94(4.43)*	70.56(6.39)
3. Condition 2 plus 20 db	75.06(3.41)*	73.26(4.29)*	71.31(5.68)
4. Condition 3 plus 20 db	78.38(4.40)*	75.78(4.95)*	72.25(5.76)*
5. Condition 4 plus 25 db	83.98(4.95)	81.14(5.71)	75.56(5.49)

* Significantly different (t) from the mean immediately below.

Table VII

A. Summary of analysis of variance of intensity of oral response (repetitions and opposites) to stimulus words heard at five levels of intensity.

Source of variation	d.f.	Mean square
Intensity (I)	4	35.77*
Types (T)	1	7.33
Subjects (S)	24	789.51
I x T	4	2.09
I x S	96	12.81
S x T	24	8.78
S x T x I	96	4.86

* Significant (F) at 5% level of confidence (compared with I x S).

B. Relative mean intensities of the levels analyzed in A.

Level of stimulus	Relative mean (db)*
1. Threshold plus 5 db	2.4
2. Level 1 plus 6 db	2.0
3. Level 2 plus 6 db	1.7
4. Level 3 plus 6 db	.0**
5. Level 4 plus 6 db	.9

* Any difference of 1.64 significant (t) at 1% level of confidence.
** 0 - 69 db (General Radio).

Table VIII

A. Summary of analyses of variance: (1) scores of three listening panels hearing word lists that were spoken sith the Ss experiencing three different levels of sidetones; (2) the intensity of the Ss.

Source of variation	db	Mean square Intelligibility (%)	Intensity (db)
Listening panels (L)	3	5462.1	
Sidetones (Si)	3	38032.08**	3863.16**
Remainder	9	3059.86	
Phrases (P)	7		12.57
Speakers (Sp)	15		1050.55
Si x P	21		24.08
Si x Sp	45		192.11
P x Sp	105		14.97
Sp x Si x P	315		15.44

** Significant (F) at the 1% level of confidence.

B. Means and standard deviations of the values analyzed in A.

Level of sidetone	Mean Response Intelligibility (%)*	Mean Intensity (db)**
0 db	56.4	41.9
-14 db	62.9	48.0
-27 db	68.8	52;9
-38 db	76.3	54.9

* All differences between means significant (t) at the 1% level of confidence.

** Any difference between means of 3.7 significant (t) at 1% level of confidence; 2.7, 5%.

Table IX

Summary of an analysis of variance of the intensity of readings of nonsense syllables in response to stimuli of equal intensity and different loudness. Stimuli, pure tones of 10 frequencies. N, Ss, 24.

Source of variation	d.f.	Mean square
Conditions	9	13.40**
Subjects	23	63.60
Remainder	207	1.44

** Significant (F) at the 1% level of confidence.

Table X

A. Summary of analyses of variance of the rates of oral responses to heard stimuli of varying rates. Stimuli: five-syllable phrases.

		Mean square		
		Responses to stimulus conditions		
Source of variation	d.f.	one rate per list	random rates	natural rates
Rates	4	1.66**	.79**	.83**
Subjects	24	.67	.60	.66
Remainder	96	.06	.02	.02

** Significant (F) at the 1% level of confidence.

B. Mean duration (sec.) of the phrases (responses) analyzed in A.

	Duration of repetition		
Duration of stimulus	one rate per list*	random rates**	natural rates***
Rate 1 (short)	1.30	1.46	1.43
Rate 2	1.47	1.64	1.51
Rate 3	1.77	1.76	1.50
Rate 4	1.77	1.80	1.57
Rate 5	1.82	1.94	1.89

* Any difference between means of .18 significant (t) at 1% level of confidence; .14, 5%.

** Any difference between means of .10 significant (t) at 1% level of confidence; .08, 5%.

*** Same as immediately above.

Table XI

Mean duration (sec.) and intensity (db) of repetitions of phrases heard over three transmission systems at each of three intensities (1. soft; 2. medium; 3. loud) and rates (fast, medium, slow).

A. Mean duration (sec.).

	Level 1*				Level 2**				Level 3***				Mean (over-all)
	Fast	Med.	Slow	Mean	Fast	Med.	Slow	Mean	Fast	Med.	Slow	Mean	
Loudspeaker	1.15	1.35	1.45	1.31	1.25	1.37	1.48	1.37	1.33	1.49	1.63	1.48	1.39
Headphones	1.22	1.49	1.55	1.42	1.26	1.44	1.55	1.42	1.34	1.54	1.66	1.51	1.45
Direct	1.19	1.35	1.54	1.36	1.30	1.44	1.57	1.44	1.38	1.52	1.72	1.54	1.45
Mean	1.19	1.40	1.51		1.27	1.41	1.53		1.35	1.52	1.67		
Mean (over-all)		1.36				1.41				1.51			

* Any difference of .109 significant (t) at 1% level of confidence; .083, 5%.
** Any difference of .092 significant (t) at 1% level of confidence; .070, 5%.
*** Any difference of .127 significant (t) at 1% level of confidence; .097, 5%.

B. Mean intensity (db).

	Level 1*				Level 2**				Level 3***				Mean (over all)
	Fast	Med.	Slow	Mean	Fast	Med.	Slow	Mean	Fast	Med.	Slow	Mean	
Loudspeaker	31.6	32.9	32.0	32.2	28.2	29.8	28.5	28.8	40.8	41.1	40.8	40.9	34.0
Headphones	34.9	35.6	34.8	35.1	28.5	28.8	27.8	28.4	40.4	41.2	40.2	40.6	34.7
Direct	30.8	30.4	30.8	30.7	29.0	29.4	28.6	29.0	40.4	29.9	39.6	40.0	33.2
Mean	32.4	33.0	32.5		28.6	29.3	28.3		40.3	40.7	40.2		
Mean (over-all)		32.6				28.7				40.4			

* Any difference of 2.37 significant (t) at 1% level of confidence; 1.80, 5%.
** Any difference of 1.70 significant (t) at 1% level of confidence; 1.29, 5%.
*** Any difference of 1.96 significant (t) at 1% level of confidence; 1.49, 5%.

Table XII

Distribution of proportions according to the level of confidence with which the hypothesis "the precision of articulation of stimulus phrases does not affect the precision of articulation of repetitions" can be rejected. N, proportions, 120. Each proportion based on 106 judgements, by 53 judges.

Number of proportions exceeding indicated probability		Level of confidence				
		30%	10%	5%	2%	1%
+.50	P (Correct)	74	62	55	49	44
-.50	P (Incorrect)	18	10	10	7	6

Table XIII

A. Summary of analyses of variance of duration and intensity of phrases as read in eight rooms representing two sizes, shapes, and reververation times. N, speakers, 184 (23/room).

Source of variation	d.f.	Mean square duration	Mean square intensity
Room size (Si)	1	5711.0**	1710.0**
Room shape (Sh)	1	29.7	2.0
Reverberation (R)	1	1219.5**	454.0**
Si x Sh	1	25.2	
Sh x R	1	34.5	25.3
Si x R	1	1174.7	
Si x Sh x R	1	198.5	
Within groups	176	127.3	13.3

** Significant (F) at the 1% level of confidence.

B. Means of duration (sec.) and intensity (db) of phrases as spoken in each room.

	Duration*				Intensity**			
	Large		Small		Large		Small	
	Dead	Live	Dead	Live	Dead	Live	Dead	Live
Circular	1.54	1.72	1.40	1.46	70.57	75.12	76.77	79.23
Rectangular	1.52	1.75	1.42	1.37	70.83	72.50	76.77	80.67
Mean	1.53	1.74	1.41	1.42	70.70	73.81	76.77	79.95

* Any difference between two means of .17 significant (t) at the 1% level of confidence; .13, 5%.

** Any difference between two means of 2.77 significant (t) at the 1% level of confidence; 2.11, 5%.

Table XIV

A. Summary of an analysis of variance of the intensity of equated phrases read in noise (- 5 db) and quiet and in each instance with S speaking in a cubicle and in an open room.

Source of variation	d.f.	Mean square
Cubicle (C)	1	0.1
Noise (N)	1	1890.0**
Subjects (S)	23	56.6
C x N	1	.1
C x S	23	5.0
N x S	23	16.4
C x N x S	23	5.0

** Significant (F) at the 1% level of confidence.

B. Means of values analyzed in A.

	Intensity (db)	
	Noise	Quiet
Cubicle	31.0	22.2
No cubicle	31.1	22.2

Table XV

A. Summary of an analysis of variance of voice intensity measures made while a presumably live microphone was at 1, 2, 3, 4 ft. from S. N, Ss, 16.

Source of variation	d.f.	Mean square rate
Subjects	15	18.59
Distances	3	1.93**
Remainder	45	.47

**Significant (F) at the 1% level of confidence.

B. Means of relative intensity values analyzed in A.

Distance from microphone (ft.)	Intensity (db)*
1	21.11
2	20.66
3	20.54
4	21.26

*Any difference between means of .67 significant (t) at the 1% level of confidence; .51, 5%.

Table XVI

Mean relative intensity of officers and officer candidates when reading phrases: (1) alone, (2) to members of their group, and (3) to members of the 'other' group.

Audience	Means of relative intensity (db)	
	Officers	Officer candidates*
Alone	39.0	34.8
To fellows	38.6	37.0
To 'other' group	38.3	36.9

*Any difference between means of 1.6 significant (t) at the 1% level of confidence.

Table XVII

A. Summary of analyses of variance of (1) the intelligibility scores assigned each of 144 Ss who were speaking under six different directions (eleven listeners/S), and (2) the mean intensity with which Ss spoke.

Source of variation	d.f.	Mean square	
		Intelligibility	Intensity
Tests	23	6.3	6.3
Directions	5	6.9	76.4**
Remainder	115	4.5	15.3

**Significant (F) at the 1% level of confidence.

B. Means of the scores analyzed in A.

Condition	Intelligibility (%)	Intensity (db)*
1. "...just short of shouting"	62	29.9
2. "...the strain of shouting"	64	32.2
3. "...good, clear sidetone"	64	28.9
4. "...(none)"	57	27.2
5. "...('say again; talk louder')"	61	28.8
6. "...you must speak loudly"	62	31.0

*Any difference between means of 3.0 significant (t) at the 1% level of confidence; 2.3, 5%.

References (also bibliography of technical reports)		No. Technical Report
1. Lightfoot, Charles	Rate of speaking: 1. relationship between original and repeated phrases.	SDC 411-1-1
2. Black, John W.	Loudness of speaking: The effect of heard stimuli on spoken responses.	SDC 411-1-2
3. Lightfoot, Charles Black, John W.	Rates of speaking: reports II-III (Responses to heard stimuli).	SDC 411-1-3
4. Lightfoot, Charles Morrill, Scott, N.	Loudness of speaking: The effect of the intensity of sidetone upon the intensity of the speaker.	SDC 411-1-4
5. Black, John W.	Timed phrases.	SDC 411-1-5
6. Black, John W.	Exploratory Studies: The relation between 'uncertainty' and vocal intensity.	SDC 411-1-6
7. Black, John W. Lightfoot, Charles Mitchell, John	The effects on vocal frequency and intensity of hearing sustained tones while reading.	SDC 411-1-7
8. Lightfoot, Charles	Some effects of the common cold.	SDC 411-1-8
9. Lightfoot, Charles	Rate of speaking: IV effects of the mode and level of transmitting messages upon the relationship between their duration and the duration of their repetition.	SDC 411-1-9
10. Lightfoot, Charles	Effects of the mode and rate of transmitting messages upon the relationship between their intensity and the intensity of repetition of them.	SDC 411-1-10
11. Black, John W.	Natural frequency, duration and intensity of vowels in reading.	SDC 411-1-11
12. Black, John W.	Inflection of repeated messages.	SDC 411-1-12
13. Black, John W.	The relation between messages types and vocal rate and intensity.	SDC 411-1-13
14. Black, John W.	The effect of room characteristics upon vocal intensity and rate.	SDC 411-1-14
15. Black, John W.	Precision of articulation in repeated phrases.	SDC 411-1-15
16. Black, John W.	A compensatory effect in vocal responses to stimuli of low intensity.	SDC 411-1-16