DOCUMENT RESUME

ED 099 932 CS 500 927

AUTHOR Sticht, Thomas G.
TITLE Some Interactions of Speech Rate, Signal Distortion, and Certain Linguistic Factors in Listening Comprehension. Professional Paper No. 39-68.
INSTITUTION George Washington Univ., Alexandria, Va. Human Resources Research Office.
REPORT NO HumRRO-PP-39-68
PUB DATE Nov 68
NOTE 13p.

EDRS PRICE MF-$0.75 HC-$1.50 PLUS POSTAGE
DESCRIPTORS Armed Forces; *Educational Research; Higher Education; *Linguistic Patterns; *Listening Comprehension; Listening Skills; *Speech; *Speech Compression
IDENTIFIERS *Signal Distortion

ABSTRACT

This experiment was designed to determine the relative effects of speech rate and signal distortion due to the time-compression process on listening comprehension. In addition, linguistic factors--including sequencing of random words into story form, and inflection and phraseology--were qualitatively considered for their effects on listening comprehension. The subjects were 180 Army inductees who had scored 80 or above on the "Armed Forces Qualification Test" (AFQT). The results indicate that both speech rate and signal distortion may affect listening comprehension, especially in the case of signal distortion when low-redundancy materials are used. (RB)

Professional Paper 39-68
December 1968

Some Interactions of Speech Rate, Signal Distortion, and Certain Linguistic Factors in Listening Comprehension

by

Thomas G. Sticht

Paper for
Psychonomic Society Meeting
St. Louis, Missouri
November 1968

HumRRO

The George Washington University
HUMAN RESOURCES RESEARCH OFFICE

The Human Resources Research Office is a nongovernmental agency of The George Washington University. The research reported in this Professional Paper was conducted under contract with the Department of the Army (DA 44-188-ARO-2). HumRRO's mission for the Department of the Army is to conduct research in the fields of training, motivation, and leadership.

The George Washington University • Human Resources Research Office
300 North Washington Street • Alexandria, Virginia 22314

Prefatory Note

This paper reports research performed under Work Unit REALISTIC, which is concerned with reading, listening, and arithmetic skills required for major military occupational specialties. The research is being conducted at Human Resources Research Office Division No. 3 (Recruit Training), Monterey, California.

SOME INTERACTIONS OF SPEECH RATE, SIGNAL DISTORTION, AND CERTAIN LINGUISTIC FACTORS IN LISTENING COMPREHENSION

Thomas G. Sticht

The rate of speech of a tape-recorded message can be accelerated by an electromechanical process. This process involves the periodic deletion of small temporal segments of the speech, while the remaining speech segments are connected to produce continuous discourse. (Foulke and Sticht, 1, give a more complete description of time compression by the sampling technique.) Through this deletion and sampling technique, the time required to present the message is decreased. Hence such speech is sometimes called time-compressed speech.

When the speech rate of materials is accelerated beyond approximately 280 words per minute (wpm) by the sampling method of compression, comprehension is found to drop at a fast rate (1). A question of concern is to what extent this drop in comprehension is due to the speech rate, as opposed to the signal distortion that results from the compression process. Previous research (2, 3, and 4) has suggested that it is the speech rate, and not the signal distortion resulting from the compression process, which produces the fast drop in comprehension. However, this suggestion is based on indirect approaches to the problem used in these studies. In none of them were the speech rate and signal distortion variables varied separately. This analytic approach was used in the present study to make it possible to evaluate the separate effects of speech rate or signal distortion on comprehension.

In prior research (4), it was suggested that the effects of signal distortion on comprehension might be attenuated by the high redundancy of the English language. (In Miller, 5, p. 106, there is a discussion of redundancy and noise and interactions thereof.) In the present study certain linguistic factors that lead to redundancy—for example, the sequential constraints imposed by syntactical rules—have been manipulated to determine some interaction effects of these factors and signal distortion on listening comprehension.

METHOD

Materials

To distinguish the effects of word distortion due to the compression process from the effects of word rate, listening selections were presented in which these parameters were varied separately. To accomplish this, phonetically balanced (PB) word lists (Egan, 6) subjected to 0, 36, and 59% compression were used.

The equipment used to prepare the compressed speech was the Eltro Information Rate Changer, distributed by Gotham Audio Corporation, 2 West 46th Street, New York, N.Y., 10036.[1] Compression is accomplished as follows:

The tape containing the recorded message passes over the surface of a cylinder that has four reproducing (playback) heads. These heads are spaced equally around the cylinder with "dead" spaces separating them. When the cylinder is stationary and the tape is moving at 15 inches per second (ips), the normal recording speed, it makes contact with one of the reproducing heads and the signal is reproduced on an auxiliary tape as recorded. To compress a message, the speed of the tape is increased and the cylinder begins to rotate in the direction of tape motion. As the speed of the tape is increased, the rotational speed of the cylinder is increased so as to maintain a relative speed of 15 ips between the tape and reproducing heads located in the cylinder. As the cylinder rotates, each of the four reproducing heads makes and then loses contact with the tape. The heads are spaced so that as one head is leaving the tape, the next is making contact, thus, reproducing a continuous message. However, since the small space between heads is dead, no message is reproduced where this dead segment comes into contact with the tape. Small segments of the recorded message are therefore discarded as the message is transferred to the auxiliary tape. The amount of compression is determined by the speed at which the tape and cylinder head move. The faster this speed, the greater the number of speech segments discarded per unit time, and the greater the amount of compression of the reproduced message. Messages can be expanded by slowing down the tape to a speed less than 15 ips. In this case, the rotation of the cylinder is reversed, and small segments of speech are reproduced. The perceptual effect is a reduction in speech rate.

From the recorded lists of phonetically balanced words, words were selected and arranged into two brief listening selections. One selection was composed of 63 words, the other of 82. To vary the speech rate of the uncompressed (0%) stories, the words were first recorded directly adjacent to one another, with no time interval between them. This process resulted in a word rate of 100 words per minute (wpm) ($\pm < 2\%$ variation) for each of the two listening selections. This was the fastest speech rate possible with the 0% compressed words. A second, slower speech rate was obtained by inserting a time interval of 180 milliseconds (msec) between words. This produced a speech rate of 75 wpm. Thus, for the two selections made up of uncompressed words, two rates of speech, 75 and 100 wpm, were available.

Listening selections composed of the words compressed by 36% were prepared in the same manner as the selections written with the uncompressed words. When recorded with no time interval between words, the

[1]Citation of this trade name and product in this paper is for research information purposes only and does not constitute official Department of the Army endorsement or approval.

speech rate was 155 wpm. By introducing temporal intervals of 222 and 402 msec between words, speech rates of 100 and 75 wpm were obtained. Using similar techniques, the listening selections were recorded with the words compressed by 59%. When recorded with no temporal interval between words, the speech rate was 222 wpm. By inserting time intervals of 111, 333, and 513 msec between words, versions of the listening selection having speech rates of 155, 100, and 75 wpm were prepared.

Through the procedures outlined above, nine different versions of the same two listening selections were obtained. In two versions, the words were not compressed, and speech rates of 75 and 100 wpm were produced. In three versions, the words were compressed by 36%, and word rates of 155, 100, and 75 wpm were produced. In four versions, the words were compressed by 59%, and rates of speech of 222, 155, 100, and 75 wpm were obtained.

To compare the effects of distortion due to compression, while holding speech rate constant, comprehension of the selections presented at either 75 or 100 wpm, with compression ratios increasing from 0 to 59%, was tested. To compare the effects of speech rate, with distortion due to compression either omitted or held constant, changes in comprehension of the passages within a category of percent compression (for example, the two versions with noncompressed words) were determined.

Because the words used in composing the two listening selections were individually pronounced, the selections contained no inflectional cues of emphasis or meaning. There were also no cues of transition from one phrase or sentence to another. In short, linguistic cues to meaning were restricted, in each selection, primarily to cues from the sequential constraints of the language, and the lexical meaning of the words.

The use of individually spoken phonetically balanced words makes possible the isolation of compression and rate factors, relatively independent of other linguistic and semantic variables. However, such speech is not "normal". To determine what the effects on comprehension might be of removing inflectional cues, pauses between sentences, and so forth, additional versions of the listening selections were recorded using a natural, prose oral reading style in which inflectional and phrasing cues were maintained. In this case, the two listening selections were recorded at an average rate of 158 wpm.

A final listening test was prepared with the same phonetically balanced words used in each of the two selections, but presented in random sequence. This was to determine what effects removal of the sequential constraint from the words in story form might have on the comprehension test. With the random PB words, the sequential constraints of the language were removed, leaving only the lexical meaning of the words.

To test comprehension, the cloze (Taylor, 7) technique was used. Typed copies of the two listening selections were prepared with every fourth word omitted. A standard length line was inserted in place of

the word. If the fourth word was at the beginning of a sentence, the next word was deleted. Conjunctions such as "and" were not deleted, so as to increase the difficulty of the tests. In the cloze test, the subjects were asked to fill in the missing words as they read the passage. The test was scored by counting the number of correct words. Only the words actually deleted were accepted as correct responses, that is, no synonyms were accepted. To find out how well the subjects could perform on the tests without prior exposure to the listening selections, a group who had not heard the selections (100% compression) read the tests and attempted to guess what the missing words were. This procedure provided baseline data to determine improvement due to listening to the stories.

All told, there were 12 test conditions, including the baseline tests. These conditions are listed in Table 1.

Table 1

Relative Effects of Signal Distortion, Speech Rate, and Linguistic Factors on Comprehension of Listening Selections

($N = 15$ per Test Condition)

Test Condition	Number Correct		Percent Correct[a]	
	Mean	SD		
Baseline Test[b]	7.07	2.15	21.4	
Random Words	10.93	1.76	33.1	*
Word Rate of 75 wpm				
0% compression	14.33	5.31	43.4	
35% compression	11.07	3.20	33.5	*
59% compression	10.00	2.45	30.3	*
Word Rate of 100 wpm				
0% compression	10.13	3.25	30.7	*
36% compression	10.07	2.40	30.5	*
59% compression	11.00	3.46	33.3	*
Word Rate of 155 wpm				
36% compression	10.80	4.72	32.7	*
59% compression	11.47	2.59	34.7	*
Word Rate of 222 wpm				
59% compression	10.60	3.16	32.1	*
Normal Prose (158 wpm)	16.00	3.46	48.4	#

[a] * indicates scores significantly different from both the 0% – 75 wpm and Normal Prose conditions; # indicates scores not significantly different from 0% – 75 wpm condition.

[b] All test scores were significantly different from the baseline test score.

Subjects

One-hundred-eighty Army inductees served as subjects, 15 in each of the 12 test conditions. All had AFQT (Armed Forces Qualification Test) scores of 80 or above. There were no significant differences between the mean AFQT scores of the individuals in the 12 conditions.

Procedure

The subjects in each experimental group were tested together in a classroom. They were asked to listen to two listening selections, and were told that following each selection they would be tested on how well they remembered what they heard. Subjects who listened to the stories composed of phonetically balanced words were told that the selections were prepared in a special way and would not sound like regular speech, so they would have to listen carefully. Following the presentation of each selection, the cloze test was administered. Subjects were instructed to guess if they did not know the answers.

The subjects who listened to the random words were told that they would hear a list of words and would then be administered a test in which they were to use the words. They were instructed to listen carefully and to remember as many words as they could. Following the presentation of the PB words used in the first listening selection, the appropriate cloze test was administered. This procedure was then repeated using those phonetically balanced words that made up the second listening selection.

The baseline test performance data were obtained in the same classroom as the experimental data and in the manner outlined above.

Presentation of all listening materials was by means of a tape recorder. The loudness was adjusted to a "comfortable" listening level by those conducting the experiment. All subjects indicated they could hear the messages satisfactorily.

RESULTS AND DISCUSSION

The results are summarized in Table 1. In all cases, those groups that listened to the stories or random words performed significantly better (median test, Siegel, 8, p. 111) on the cloze test than those who took the test without prior exposure to the missing words in the listening selections. Thus, the listening experiences did improve test performance.

Speech Rate and Signal Distortion

Regarding the speech rate and compression distortion variables, performance was significantly better for the 0%—75 wpm condition than for all other combinations of these variables. There were no significant differences between the scores for any of the remaining speech rate and stimulus compression combinations.

Holding compression constant, speech rate effects were found only with the 0% compressed selections when the rate was increased from 75 to 100 wpm. This increase produced a significant reduction in test scores. With the individually pronounced phonetically balanced words it was not possible to produce the very fast word rates which have resulted in large decrements in comprehension in other experiments (1). The limitation on the rate of speech in the present study was due to the fact that the phonetically balanced words averaged approximately .6 sec in duration when spoken individually. Thus, the fastest word rate possible, with no temporal interval between words, was about 100 wpm. Even when the words were compressed by approximately 59%, and abutted without any interval between them, a top rate of only 222 wpm was obtained. This rate is well within the range which would not be expected to affect comprehension by any significant amount. Quite possibly, faster word rates might reduce comprehension below the levels observed in the present study.

The effects of signal distortion due to compression are indicated by the decrease in scores in the 75 wpm condition where the compression was increased from 0% to 36%. The additional compression to 59% produced a slight additional decrease in test performance. These results contrast with those of prior work (4) from which it was concluded that signal distortion did not significantly reduce comprehension. At that time it was suggested that signal distortion might be a more potent factor with materials of reduced redundancy. The present study used low-redundancy materials. Such aids to efficient message encoding as inflection and phraseology were omitted. In addition, the use of PB words produced many low-probability sequences in the messages. The results obtained with this low-redundancy material appear to reinforce the suggestion that signal distortion will be a more or less potent factor for listening comprehension depending on the redundancy of the listening selection.

Further Observations on Linguistic Factors

The material used to study the effects of speech rate and signal distortion in the present study differed from normal connected discourse. In these materials there were no cues to meaning in the form of inflectional, phrasing, or transitional cues between sentences. There were, however, cues of sequential dependency (albeit reduced, since the subjects could not be certain at times where contingencies began and ended) and lexical cues.

The importance of these various linguistic factors for comprehension test performance is illustrated in Figure 1. It is shown in this figure that listening comprehension improves when cues to meaning are increased from none (listening selection not heard) to include lexical, syntactical, and inflectional-phrasing cues. Thus, it is seen that the successive addition of these various cues is associated with improved performance. An additional factor to be considered with the uncompressed, sequenced words and the normal prose passage is the speech rate. For the former, the speech rate was 75 wpm, while for

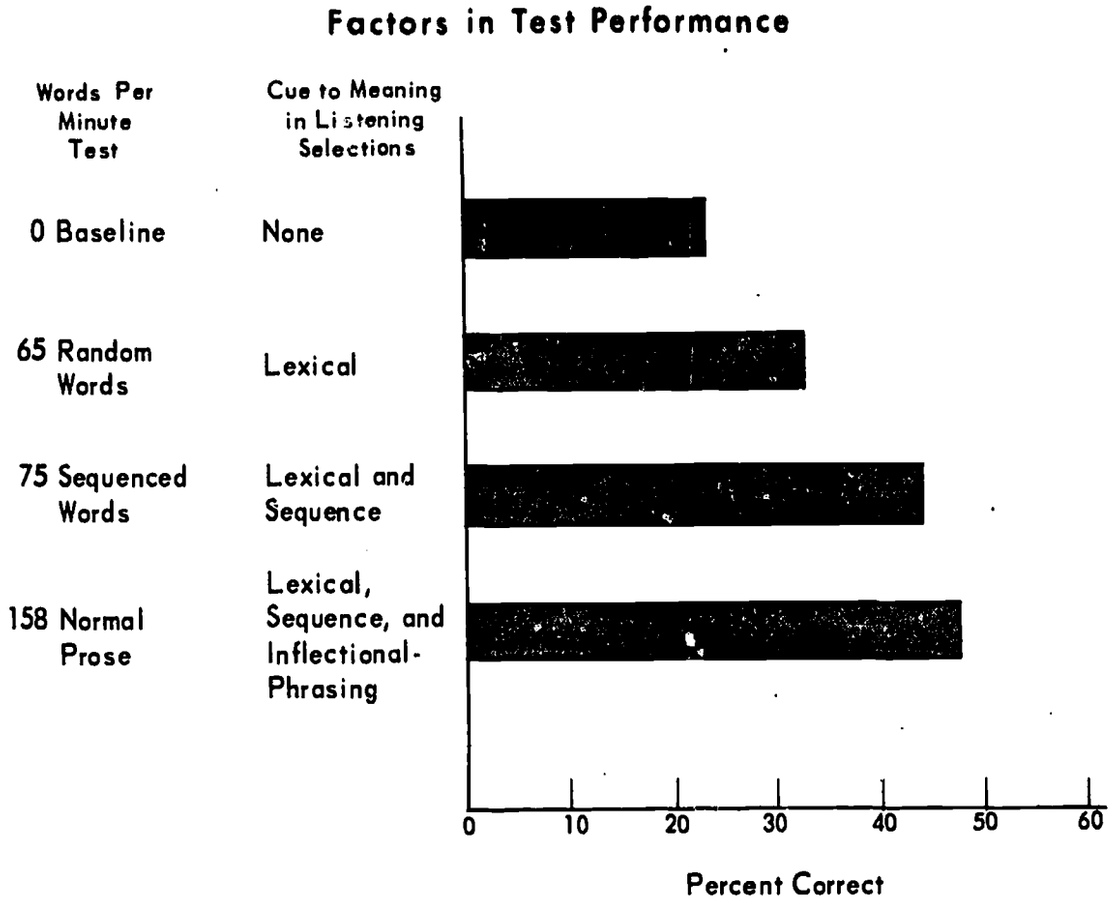

Figure 1

the normal prose, the speech rate was 158 wpm. A speech rate of 158 wpm falls within a "normal" range of 90—175 wpm, whereas the 75 wpm rate lies outside this range. Hence, the comprehension of the normal prose may have been aided by a more normal time frame, as well as the additional expressional linguistic factors. It should be mentioned, too, that while the comprehension scores for the sequenced words and normal prose are not significantly different, it is the general trend shown in Figure 1 that is of concern. Because of the differences in word rate of the three listening conditions, Figure 1 should be considered simply as a qualitative indication of the importance of the various linguistic factors. A more rigorous demonstration would require similar word rates.

Performance, Perception, and Storage/Retrieval Processes

In the present study equal performance was obtained with the randomly presented words and the compressed words presented in story form at 75 wpm. The latter performance was significantly below that obtained with the uncompressed stories presented at 75 wpm. A question of interest is whether the reduction in comprehension indexed with the compressed stories was due to misperception, as commonly found with individually presented words in intelligibility test, or

to storage/retrieval problems (with the possibility of non-perception included as a storage problem).

The answer to this question is suggested by a comparison of the error responses obtained with the uncompressed, randomly presented words and the compressed stories presented at 75 wpm. Unpublished observations of the error responses in intelligibility tests using compressed phonetically balanced words indicate that these errors almost always involve a phonemic substitution in the word such that it rhymes with the target word, for example, such as substituting lay for way. Thus, if misperceptions were the primary cause for the decrease in performance observed with the compressed materials in the 75 wpm condition of the present study, it seems likely that a large number of the error responses would rhyme with the correct responses. If misperception was not the primary factor operating to reduce the comprehension of these compressed materials, then the error responses should equal the heterogeneity of error responses obtained with the uncompressed, randomly presented words, and would not necessarily rhyme with the target word.

An analysis of error responses as suggested above showed that the error responses obtained with the compressed materials were as heterogeneous as those obtained with the randomly presented words. This suggests that storage/retrieval problems, and not misperception, were likely to be instrumental in reducing performance with the compressed materials.

The present data shed no light on the nature of the storage/retrieval factors which may have reduced performance with the compressed materials. However, the data are relevant to previous research concerned with speech rate, word duration, and storage/retrieval problems. Arronson (9) conducted studies in which the duration of spoken digits was varied, while the interdigit interval was varied to maintain a presentation rate of three per second. Seven-digit sequences were presented to subjects who attempted to memorize each sequence and recall it a few seconds after the presentation. She found that recall accuracy was higher for the briefer stimulus durations with the larger interdigit intervals between them. For a fixed presentation rate, the subjects produced fewer errors when the ratio of speech to silent time between digits was decreased. Arronson suggested that the silent time between stimuli may be used for perceptual processing, and hence may be more important than the stimulus duration, at least within some finite limit of the latter.

In the present study the duration of spoken words was varied while the interword interval was changed to keep the presentation rate constant. This was similar to Arronson's procedure with digits. However, in the present study recall did not improve when word duration was decreased and the time between words was increased. In fact, the opposite occurred, as indicated by the data in Table 1 obtained with the 75 wpm speech rate and 0%, 36%, and 59% compression ratios. This indicates that Arronson's hypothesis must be restricted to the materials and procedures of her experiments; at least, it is not confirmed by the present study, in which procedures more nearly approximate the usual listening comprehension test situation.

SUMMARY

This experiment was designed to determine the relative effects of speech rate and signal distortion due to the time-compression process on listening comprehension. In addition, linguistic factors, including sequencing of random words into story form, and inflection and phraseology were qualitatively considered for their effects on listening comprehension.

The results indicate that both speech rate and signal distortion may affect listening comprehension. The latter effects became noticeable with the low-redundancy material used in the present study, while they were not apparent with the more redundant material of prior work done by the author (4). The effects of signal distortion in the present study did not appear to result from misperception of the distorted words. The addition of sequencing (syntax) and inflectional and phraseology cues to meaning in the materials used in this study tended to improve listening comprehension.

LITERATURE CITED

1. Foulke, E. and Sticht, T.G. "A Review of Research on Time Compressed Speech," in *Proceedings of the Louisville Conference on Time Compressed Speech, October 19-21, 1966*, E. Foulke (ed.), Center for Rate Controlled Recordings, University of Louisville, Louisville, Ky., 1967.

2. Foulke, E. and Sticht, T.G. "The Intelligibility and Comprehension of Time Compressed Speech," in *Proceedings of the Louisville Conference on Compressed Speech, October 19-21, 1966*, E. Foulke (ed.), Center for Rate Controlled Recordings, University of Louisville, Louisville, Ky., 1967.

3. Foulke, E. "The Perception of Time Compressed Speech," paper for The Perception of Language Conference, University of Pittsburgh, Pittsburgh, Pa., January 1968.

4. Sticht, T.G. "Signal Distortion Versus Speech Rate in the Comprehension of Time-Compressed Speech," manuscript in preparation, 1968.

5. Miller, G.A. *Language and Communication*, McGraw-Hill, New York, 1951.

6. Egan, J.P. "Articulation Testing Methods," *Laryngoscope*, vol. 58, 1948, pp. 955-991.

7. Taylor, W. "Cloze Procedure: A New Tool for Measuring Readability," *Journalism Quarterly*, vol. 30, 1953, pp. 415-433.

8. Siegel, Sidney. *Nonparametric Statistics for the Behavioral Sciences*, McGraw-Hill, New York, 1956.

9. Arronson, Doris. "Temporal Factors in Perception and Short-term Memory," *Psychol. Bull.*, vol. 67, 1967, pp. 130-144.