DOCUMENT RESUME

ED 066 080 FL 003 355

AUTHOR Sticht, Thomas G.
TITLE Mental Aptitude and Comprehension of Time-Compressed and Compressed-Expanded Listening Selections.
INSTITUTION Human Resources Research Organization, Alexandria, Va.
SPONS AGENCY Office of the Chief of Research and Development (Army), Washington, D.C.
REPORT NO HumRRO-PP-6-72
PUB DATE Mar 72
NOTE 11p.; Reprint from Journal of Auditory Research; v10 p103-109 1970

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS Articulation (Speech); Audiometric Tests; Audiometrists; Comprehension; Electromechanical Technology; *Experiments; *Information Processing; *Language Research; *Listening Comprehension; Machine Translation; Phonemes; Phonology; Psycholinguistics; Speech; *Speech Compression

ABSTRACT

The comprehensibility of materials compressed and then expanded by means of an electromechanical process was tested with 280 Army inductees divided into groups of high and low mental aptitude. Three short listening selections relating to military activities were subjected to compression and compression-expansion to produce seven versions. Data indicate that expanding previously compressed materials to restore the word rate to normal may restore the comprehension of the material to very near normal when the compression/expansion is limited to 40%. Present results substantiate findings that factors limiting the comprehensibility of rapid speech reside more with the inability of the listener to process rapid rates of speech than with the signal distortion produced by the equipment or compression process. (Author)

DOD
FL

HumRRO

Professional Paper 6-72
HumRRO-PP-6-72

Mental Aptitude and Comprehension of Time-Compressed and Compressed-Expanded Listening Selections

Thomas G. Sticht

HUMAN RESOURCES RESEARCH ORGANIZATION
300 North Washington Street • Alexandria, Virginia 22314

Reprinted from
Journal of Auditory Research
Volume 10, 1970, pp. 103-109

March 1972

Prepared for
Office of the Chief of Research and Development
Department of the Army
Washington, D.C. 20310

The Human Resources Research Organization (HumRRO) is a nonprofit corporation established in 1969 to conduct research in the field of training and education. It is a continuation of The George Washington University Human Resources Research Office. HumRRO's general purpose is to improve human performance, particularly in organizational settings, through behavioral and social science research, development, and consultation. HumRRO's mission in work performed under contract with the Department of the Army is to conduct research in the fields of training, motivation and leadership.

Published
March 1972
by
HUMAN RESOURCES RESEARCH ORGANIZATION
300 North Washington Street
Alexandria, Virginia 22314

BIBLIOGRAPHIC DATA SHEET	1. Report No. HumRRO-PP-6-72	2.	3. Recipient's Accession No.
4. Title and Subtitle MENTAL APTITUDE AND COMPREHENSION OF TIME-COMPRESSED AND COMPRESSED-EXPANDED LISTENING SELECTIONS			5. Report Date Mar 72 6.
7. Author(s) Thomas G. Sticht			8. Performing Organization Rept. No. PP 6-72
9. Performing Organization Name and Address Human Resources Research Organization (HumRRO) 300 North Washington Street Alexandria, Virginia 22314			10. Project/Task/Work Unit No. REALISTIC 11. Contract/Grant No. DAHC 19-70-C-0012
12. Sponsoring Organization Name and Address Office, Chief of Research and Development Department of the Army Washington, D.C. 20310			13. Type of Report & Period Covered Professional Paper 14.

15. Supplementary Notes HumRRO Division No. 3, Presidio of Monterey, California; reprinted from Journal of Auditory Research, 1970, vol. 10, pp. 103-109.

16. Abstracts

The comprehensibility of materials compressed and then expanded by means of an electromechanical process was tested with 280 Army inductees divided into two groups of high- and low-mental aptitude. Three short listening selections relating to military activities were subjected to compression and compression-expansion to produce seven versions. Data indicate that expanding previously compressed materials to restore the word rate to normal may restore the comprehension of the material to very near normal when the compression/expansion is limited to 40%. Present results substantiate findings that factors limiting the comprehensibility of rapid speech reside more with the inability of the listener to process rapid rates of speech than with the signal distortion produced by the equipment or compression process.

17. Key Words and Document Analysis. 17a. Descriptors

*Machine translation
*Phonemes
*Phonology
*Psycholinguistics
*Speech
*Speech analysis
*Speech articulation
*Audiometry
*Comprehension

17b. Identifiers/Open-Ended Terms

Listening
Speech rate
Signal distortion
Time-compressed speech
Recorded speech transmissions

17c. COSATI Field/Group 0507 Behavioral and social sciences/Linguistics
0510 Behavioral and social sciences/Psychology

18. Availability Statement	19. Security Class (This Report)	21. No. of Pages
Approved for public release; distribution unlimited.	UNCLASSIFIED	11
	20. Security Class (This Page) UNCLASSIFIED	22. Price

FORM NTIS-38 (10-70) USCOMM-DC 40329-P71

Prefatory Note

The research reported in this paper was performed under Work Unit REALISTIC, Determination of Reading, Listening, and Arithmetic Skills Required for Major Military Occupational Specialties, at the Human Resources Research Organization Division No. 3, Presidio of Monterey, California. Dr. Sticht is a Senior Staff Scientist with the HumRRO Division.

Dr. Sticht is the author of other HumRRO publications directly related to the area of the REALISTIC research reported in this paper. They include *Learning by Listening in Relation to Aptitude, Reading, and Rate-Controlled Speech*, Technical Report 69-23, December 1969; *Studies on the Efficiency of Learning by Listening to Time-Compressed Speech*, Professional Paper 4-70, February 1970; *Comprehension of Repeated Time-Compressed Recordings*, Professional Paper 2-70, January 1970; *Some Interactions of Speech Rate, Signal Distortion, and Certain Linguistic Factors in Listening Comprehension*, Professional Paper 39-68, December 1968; and *Some Relationships of Mental Aptitude, Reading Ability, and Listening Ability Using Normal and Time-Compressed Speech*, Professional Paper 33-68, October 1968.

Journal of Auditory Research, 1970, 10, 103-109

MENTAL APTITUDE AND COMPREHENSION OF TIME-COMPRESSED AND COMPRESSED-EXPANDED LISTENING SELECTIONS

THOMAS G. STICHT
Human Resources Research Organization[1]
Monterey, California

INTRODUCTION

By means of an electromechanical process, recorded speech can be accelerated or decelerated without accompanying changes in the frequency spectrum, as typically occurs whenever a recording is played at some rate faster or slower than the original recording rate. The electromechanical process for accelerating speech is analogous to cutting out and discarding small, periodic samples of a tape, and splicing the remainder together to form a continuous tape. This process depends upon the fact that the duration of most speech elements of phonemes is greater than actually needed for perception of the speech sounds. Due to this temporal redundancy, a considerable portion (up to 75% in some cases) of a word may be deleted without totally imparing its intelligibility (Foulke & Sticht, 1969, p. 53). Because the acceleration process reduces the amount of time required to present a message, the message is said to be time-compressed.

Time-expanded speech is produced by periodically repeating a small segment of a recorded message. This produces a perceptual deceleration of the speech so that it sounds slower than the normal recorded speech. By combining speech compression with speech expansion a recorded message may be compressed for rapid transmission over a crowded channel and can then be expanded at the destination to restore the speech rate. Several studies (cf., Foulke & Sticht, 1969) have explored the effects of speech acceleration upon the comprehension of recorded messages. A typical finding is that speech may be accelerated up to around 275 wpm (words-per-minute) without seriously imparing the comprehensibility of the message.

Less is known about the comprehensibility of materials which have been compressed and then expanded. In one study (Sticht, 1969) it was found that, whereas accelerating speech to 275 wpm (40% compression) produced a significant decrease in the comprehensibility of the message, restoring the speech rates to normal by the expansion of the compressed materials restored the comprehensibility of the message to normal.

The present research extended the foregoing analyses to include the compression of speech by 20% (206 wpm), 40% (275 wpm), and 47% (300 wpm) with expansion to normal. In addition, two groups of Ss were used, of high and low mental aptitude. The research cited above involved only high-aptitude Ss. However, previous research has indicated that the listening skills of low-aptitude men differ considerably from those of

[1]The research reported in this paper was performed at HumRRO Division No. 3, Monterey, California, under Department of the Army contract with The Human Resources Research Organization; the contents of this paper do not necessariiy reflect official opinions of policies of the Department of the Army. Reproduction in whole or in part is permitted for any purpose of the Department of the Army.

high-aptitude men (Sticht, 1968). Therefore it was desirable in the present study to determine if the compression/expansion process would produce comparable results for both levels of aptitude.

METHOD

Subjects. The Ss were 280 Army inductees from Ft. Ord, California. Half of the men had Armed Forces Qualification Test (AFQT) scores of 80 or above, (Hi-aptitude) and half had AFQT scores of 30 or below (Lo-aptitude). These mental aptitude test scores are not intelligence test scores. Rather, they indicate ability to benefit from military training and are the resultant of both heredity and educational experience. In terms of intelligence test scores, an AFQT score of 80 corresponds roughly to a Wechsler I.Q. of 110-115, while an AFQT score of 30 would correspond roughly to a Wechsler score of 80-90 (Hedlund, 1959).

The 140 men in each aptitude group were divided into 7 subgroups, each containing 20 men. Each group listened to a different version of three recorded messages. Group 1 listened to the recordings presented uncompressed at a normal speech rate of 165 wpm. Groups 2, 3, and 4 listened to the same selections presented at compression ratios of 20% (206 wpm), 40% (275 wpm), and 47% (300 wpm). Groups 5, 6, and 7 listened to the same tapes as heard by Groups 2, 3, and 4, but in each case the compressed tapes were expanded to restore the speech rate to normal (165 wpm). Thus, Groups 5, 6, and 7 listened to tapes that were first compressed and then expanded.

Materials. Three listening selections were prepared, each concerning some activity related to military service. The first selection concerned a combat situation, the second presented fire drill instruction, and the third selection described the transfer unit of a 2½ ton truck. The time required to listen to each of the selections in uncompressed form was 55 sec, 36 sec, and 56 sec respectively.

The listening selections were subjected to compression and compression/expansion to produce the seven versions described above. Compression and expansion were accomplished by means of the Eltro Information Rate Changer.[2]

Procedure. The seven groups were tested on different days in an ordinary classroom. Hi- and Lo-aptitude Ss were tested at the same session. Ss were seated in a semi-circle about the tape recorder used to present the listening selections. The Ss were told that they were going to be tested to determine how well they could remember some listening selections. They were told that they would hear three listening selections, and that following each selection they would be asked questions about the selection. They were instructed to write their answers on the answer sheets provided. Questions from the Ss were answered, and the listening selections and comprehension tests were administered. For each selection, the comprehension tests were administered. For each selection, the comprehension test included 12 "fill-in-the-blank" questions about factual information in the selection. The questions were read aloud and were repeated as often as requested. The same procedure was followed for the uncompressed, compressed, and compressed/expanded versions of the listening selections. All materials were presented at a "comfortable" listening level established by the Ss.

[2]Equipment is identified for purposes of documentation and does not imply endorsement by either Human Resources Research Organization or Department of the Army.

RESULTS AND DISCUSSION

Table I presents the means and standard deviations of the comprehension scores in terms of the number of questions correctly answered. The scores are summed for all three listening selections. Thus, the maximum score possible is 36 correct. Table I also presents biographical data for the various groups.

Fig. 1 presents the data from Table I in graphic form, and transformed into percent correct scores. Analysis of variance was performed for the three pairs of groups who listened to compressed and compressed/expanded messages (thus the unpaired, uncompressed conditions were omitted from the analysis). Table II summarizes the analysis of variance. In this analysis the B factor, compression ratio, refers to the three levels of compression (20%, 40%, 47%) used to prepare both compressed and compressed/expanded materials. The C factor, speech rate, refers to the compressed materials, in which the speech rate was increased, and the compressed/expanded materials in which the speech rate was constant at 165 wpm.

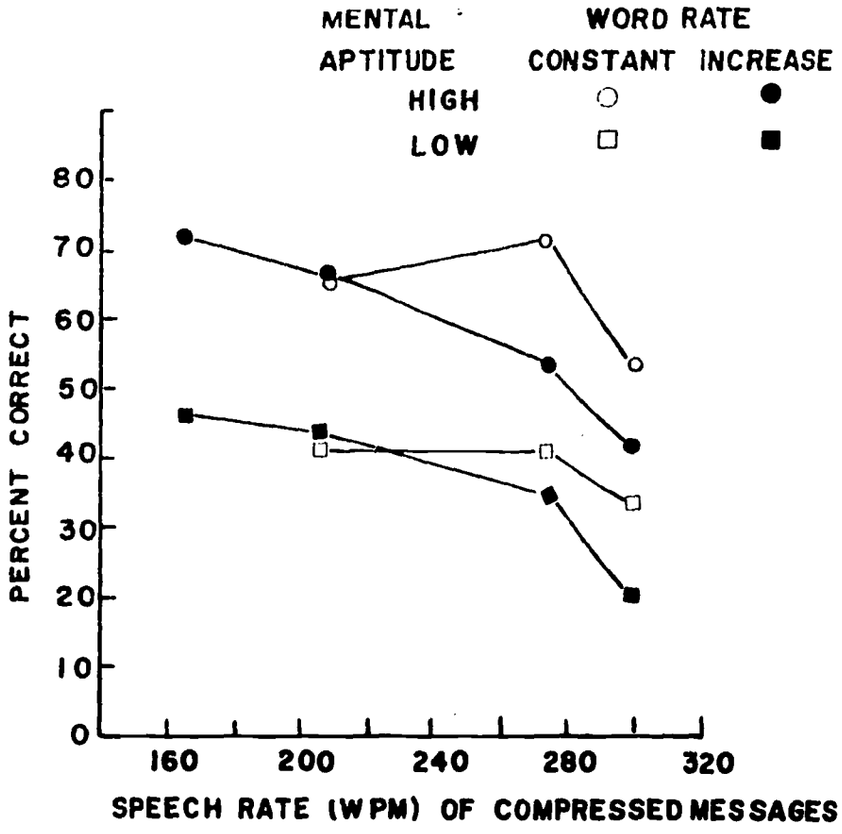

Fig. 1: Per cent correct comprehension scores as a function of speech rate.

Table I. Means and standard deviations of the comprehension scores for high and low mental aptitude Ss who listened to either normal, time-composed, or time-compressed-and-then-expanded messages.

		Age		Education		Mental Aptitude (AFQT)		Score	
Groups		Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
Hi Apt	OC[1]	21.40	2.06	14.70	1.89	88.10	5.50	25.85	4.90
Lo Apt	OC	18.00	1.30	11.18	1.04	22.00	5.80	16.75	4.30
Hi Apt	20C[2]	22.00	2.49	14.12	1.97	90.20	5.49	24.05	4.68
Lo Apt	20C	19.90	2.38	10.92	1.28	20.75	5.66	15.65	2.93
Hi Apt	20C/E[3]	20.50	2.35	13.70	2.39	88.60	5.32	23.85	3.26
Lo Apt	20C/E	17.85	1.60	10.68	1.03	24.00	5.35	14.80	4.96
Hi Apt	40C	21.15	1.69	13.75	2.00	90.25	3.86	19.45	4.20
Lo Apt	40C	18.70	1.92	10.65	1.39	21.35	5.20	12.70	4.35
Hi Apt	40C/E	21.00	2.64	14.05	2.37	89.25	6.02	26.05	3.44
Lo Apt	40C/E	18.75	1.83	10.55	1.22	23.80	4.19	14.80	4.50
Hi Apt	47C	20.60	1.10	13.60	1.70	87.20	5.84	15.05	4.07
Lo Apt	47C	19.65	1.09	11.20	1.40	22.40	5.53	7.65	2.83
Hi Apt	47C/E	21.45	1.91	14.60	2.01	90.80	6.13	19.45	3.91
Lo Apt	47C/E	19.70	3.79	11.50	1.19	22.05	5.99	12.15	4.02

[1]OC = 0% Compression; [2]20C = 20% Compression; [3]20C/E = 20% Compression/Expansion

The significant interaction of compression ratio and speech rate (BC) is indicated in Fig. 1 by the divergence of the curves for which the speech rate was constant from the curves for which speech rate was increased. Tests of the simple effects of increased vs. constant speech rate at each of the three compression levels indicated no significant differences between the 20% compressed (206 wpm) condition and the 20% compressed/expanded (165 wpm) condition. The remaining two pairs differed significantly ($p < .001$). A separate analysis of variance was performed on the data for the two aptitude groups for the uncompressed, 20% compressed and 20% compressed/expanded

Table 11. Analysis of Variance: Comprehension of Compressed vs. Compressed/Expanded Materials

Source	df	MS	F	P
Aptitude (A)	1	4,191.70	263.96	< .001
Compression Ratio (B)	2	797.26	50.21	< .001
Speech Rate (C)	1	456.50	28.75	< .001
AB	2	15.64		
AC	1	42.51		
BC	2	161.76	10.19	< .005
ABC	2	30.45		
Within Cell	228	15.88		
Total	239			

conditions and indicated no significant differences due to those conditions (aptitude was significant, $p < .001$).

These data indicate that, using the present equipment, expanding previously compressed materials to restore the word rate to normal may restore the comprehension of the material to very near normal–when the compression/expansion is limited to 40%. When the materials are compressed/expanded by 47%, there is apparently enough noise and/or signal distortion added to reduce comprehensibility of the material significantly below normal, although the restoration of a normal word rate appears to improve the comprehensibility of the material to a limited degree. These effects appear to hold for both high- and low-aptitude men.

It was previously found (Sticht, 1969) that, with high-aptitude Ss similar to those of the present study, expanding selections previously compressed by 40% to return speech rate (275 wpm) to normal (165 wpm), restored the comprehension of the material to normal. It was concluded that the reduction in comprehension with the compressed material was therefore due to the speech rate, and not the signal distortion produced by the compression process. This conclusion followed from the fact that, although the expansion process added additional signal distortion to the compressed tapes, it restored the speech rate to normal, and comprehension also returned to normal. Thus, while signal distortion was common to both the compressed and compressed/expanded tapes, the former presented materials at an accelerated speech rate, while the latter presented messages at a "normal" rate, and the comprehension improved. Those findings led to the conclusion that the factors limiting the comprehensibility of rapid speech resided more with the inability of the listener to process rapid rates of speech than with the signal distortion produced by the equipment (or compression process). The present results substantiate the previous findings and conclusions for materials compressed up to 40%. However, when materials are compressed 47% and then expanded to restore the word rate to normal, there appears to be a significant amount of signal distortion to prevent the restoration of comprehension to normal.

Because the compressed/expanded materials contain distortions and noise due to both compression and expansion, it is not clear to what extent the signal degradation accompanying the higher rates of compression alone may interact with the speech acceleration factor to produce the generally observed decrements in comprehension. However, an estimate of the degree to which the signal distortion factor may influence comprehension may be obtained by comparing the comprehensibility of materials which have been compressed by 47% with the same materials subjected to equal or greater amounts of expansion. The compression process produces distortion by periodically deleting a brief segment of the recorded speech and joining together the remaining signal segments. This brings together speech segments whose boundaries do not match exactly. In the expansion process, signal distortion is introduced by periodically repeating small segments of the speech stream. Again, this brings together speech segments with unmatched boundaries.

If the frequency of repetition in the expansion process is equal to or greater than the frequency of deletion in the compression process, similar or greater amounts of signal distortion in the form of segmental boundaries will be introduced into the expanded message as is produced in the compressed message. However, the expansion process produces a decrease in speech rate while the compression process produces an increase in the speech rate. Thus, by comparing the comprehensibility of materials expanded or compressed to produce similar frequencies of reproduction or deletion, the effects of signal distortion with and without rapid speech rates can be explored.

In Table III are presented a sub-set of data from research in progress which compares the comprehensibility of 150 word, 5th grade reading selections expanded or compressed by three different amounts. The material compressed by 58% contains more signal distortion due to repetition boundaries, and the 16% expanded materials less such distortion than was produced by the periodic deletion of speech segments in the materials compressed by 47%. Each mean is the average comprehension test score of a group of 17 High (AFQT $\geq$ 80) or Low (AFQT $\leq$ 30) aptitude Army inductees.

The data of Table III indicates that, although there was more distortion in the material expanded by 58% than in the 47% compressed materials, the latter was less comprehensible than the former. This appears to be true for both high and low aptitude men (as evaluated by t-tests, the compressed materials differed significantly from the expanded materials, while the expanded materials were not significantly different,

Table III. Comprehension Test Scores of High and Low Aptitude Ss who Listened to Time Expanded or Compressed Speech

Test Material	Hi APT		Low APT	
Expanded	Mean	S.D.	Mean	S.D.
58% - 125 wpm	80.1	2.93	58.0	11.5
16% - 175 wpm	82.6	2.97	55.0	12.2
Compressed				
47% - 375 wpm	69.6	4.25	41.2	11.0

$p < .05$). These findings suggest that at even the larger levels of compression, the speech rate factor is a more potent determiner of the comprehensibility of materials than is the signal distortion.

The foregoing conclusion is further suggested by the work of Fairbanks, et al. (1957) who showed that compressing materials by as much as 50% on their sampling equipment produced very little loss in comprehension. In their case, the uncompressed speech rate was 141 wpm and the 50% compressed rate was 282 wpm–a rate found by many to only slightly affect comprehension (Foulke and Sticht, 1969). For the data of Table III, 47% compression of a message originally recorded at 200 wpm produced a word rate of 375 wpm. Thus, although Fairbanks et al. introduced a greater number of segment boundaries, and thus more distortion than in the present study, with their materials compressed by 50%, the resulting speech rate was not sufficient to reduce comprehension to any notable degree. It appears, then, that previous conclusions (Sticht, 1969) still hold "... the barrier to the comprehension of fast rates of speech appears to be within the information processing capacities of the listener, and not in the fidelity of the time compressed signal. In short, the problem is primarily due to human, not equipment, shortcomings."

REFERENCES

1. Fairbanks, G., Guttman, N., and Miron, M.S. Effects of time compression upon the comprehension of connected speech. J. Speech Hear. Dis., 1957, 22, 10-19.
2. Foulke, E. and Sticht, T.G. Review of research on the intelligibility and comprehension of accelerated speech. Psychol. Bull., 1969, 72, 50-62.
3. Hedlund, J.L. Army classification scores: their meaning and clinical implication. Paper prepared for the Conf. Army Clin. Psychol., Washington, D.C., September, 1969.
4. Sticht, T.G. Some interactions of speech rate, signal distortion, and certain linguistic factors in listening comprehension. Audio-Visual Commun. Rev., 1969, 17, 159-171.
5. Sticht, T.G. Some relationships of mental aptitude, reading ability, and listening ability using normal and time-compressed speech. J. Comm., 1968, 18, 243-258.