By Carl Jón Denbow

Listenability and Readability:
An Experimental Investigation

Test of comprehension of written, paced written and spoken material in both easy and hard versions finds no difference between modalities.

▶ The relationship between the spoken and written word has been of interest to thinkers of philosophical bent for a long time. It has even been suggested that the nature of a whole society may depend to a great extent on whether it stresses oral or written communication.

Cultures vary greatly in their exploitation of the various senses. It has been a commonplace that the ancient Hebrews and the ancient Greeks differed in the value they placed on the auditory. The Hebrews tended to think of understanding as a kind of hearing, whereas the Greeks thought of it more as a kind of seeing, although far less exclusively as seeing than post-Cartesian Western man generally has tended to do.[1]

[1] Walter J. Ong, *The Presence of the Word* (New York: Simon and Schuster 1967), p. 3.

[2] Fang has developed a formula specifically for oral material. It is called the Easy Listening Formula (ELF). See: Irving E. Fang, "The 'Easy Listening Formula,'" *Journal of Broadcasting*, 11:63-68 (Winter 1966-67).

[3] Listenability can be thought of as a measure of "ease of listening." That is, the observed listening comprehension for material at a particular grade level as determined by some formula.

[4] George R. Klare, *The Measurement of Readability* (Ames, Iowa: Iowa State University Press, 1963), pp. 151-53.

[5] See: Robert P. Larsen and D.D. Feder, "Common and Differential Factors in Reading and Hearing Comprehension," *The Journal of Educational Psychology*, 31:241-52 (April 1940); Harry Goldstein, *Reading and Listening Comprehension at Various Controlled Rates*, Columbia University Contributions to Education No. 821 (New York: Teachers College, Columbia University, 1940); Jeanne S. Chall and Harold F. Dial, "Predicting Listener Understanding and Interest in Newscasts," *Educational Research Bulletin*, 27:141-153, 168 (Sept. 15, 1948); George R. Klare, "Understanding and Indefinite Answers to Public Opinion Questions," *International Journal of Opinion and Attitude Research*, 4:91-96 (Spring

In recent years the problem of modality differences has been increasingly approached from the more restrictive perspective of behavioral science. An aid to such studies has been the development of formulas which can be applied to a language sample to predict accurately how difficult the message is. These indices have come to be called readability formulas. As the term implies these formulas were meant primarily for gauging the comprehension of the written word.[2] Occasionally, though, they have been applied to spoken language in studies which have often been termed listenability research.[3] The results are not uniform, and have led to about as many new questions as answers. Some studies have indicated positive correlations of readability scores to listenability criteria, while others have found negative relationships, and still others have found no relationship at all.[4]

It is clear from these studies that the relationship between listenability and readability, and the corresponding relationship between listening and reading comprehension, is not simple. A few qualified generalizations, though, can be made from the literature.[5]

Listening comprehension is sometimes equivalent, and sometimes superior to reading comprehension when the material is easy. Listening comprehension is seldom found inferior with easy material.

On the other hand, when the material

▶ The author is an assistant professor of journalism at Murray State University. Material for this article came from his 1973 dissertation at Ohio University. He wishes to thank Dr. Guido H. Stempel III and Dr. George R. Klare for their help.

is difficult, reading comprehension has been found to be either equivalent to or superior to listening comprehension. Seldom has reading comprehension been shown inferior with difficult material.

Studies looking at the relationship between listening comprehension and rate of presentation seem to indicate that the latter has only negligible effect on the former between 150 wpm and 200 wpm, the upper ends of the normal speaking range. After this, though, comprehension begins to decrease more rapidly.[6]

The hearer's lack of ability to go back and re-hear a portion of a spoken message is, of course, well known to those journalists who work for broadcast outlets. Bliss and Patterson, for instance, in their undergraduate newswriting text emphasize that it is important to keep this in mind when writing broadcast copy. "The news must be written as lucidly as possible because the listener has to understand at once what he hears—there is no going back to listen again."[7]

Some of broadcast journalism education is, in fact, based on the assumption that the broadcast news story must be more readable than a newspaper story on the same subject. It has been common practice, for instance, in introductory broadcast journalism classes to ask students to rewrite traditional-style newspaper stories into broadcast style. This latter style includes the use of shorter sentences and a more conversational tone. These elements are similar to the variables used to predict readability level in many formulas. Sentence length figures in a majority of readability formulas,[8] and the desire to produce a more conversational tone would appear to be conducive to the use of more familiar words, as utilized in some readability formulas, such as the Dale-Chall, whose word lists serve to give an index of how many easy or familiar words are used in comparison to the number of hard or unfamiliar words.

Such studies, in addition to helping us find out if the broadcast journalist's assumptions are correct, will aid in defining the proper function of broadcast news and print media news. Research indicates that the informed citizen uses both, but we do not know if the effectiveness has been maximized or not.

Method

Experimental Design. In order to assess these assumptions a pre-test/post-test experimental design consisting of a total of eight groups—three experimental groups and one control group, at each of two readability levels—was employed. The treatments received by each group are:

Normal Reading Group (NR)—*S*s re-

1950); Robert C. Nuckols, "Verbi," *International Journal of Opinion and Attitude Research,* 3:575-586 (Winter, 1949-50); James R. Young, "Understanding Radio News: The Effect of Style," JOURNALISM QUARTERLY, 27:19-23 (March, 1950) Francis A. Cartier, "An Experimental Study of the Effect of 'Human Interest' Factors on Listenability" (Unpublished Doctor's dissertation, University of Southern California, 1951); K.C. Beighley, "An Experimental Study of the Effect of Four Speech Variables on Listener Comprehension," *Speech Monographs* 19:249-258 (Nov. 1952); William Allen, "Readability of Instructional Film Commentary," *The Journal of Applied Psychology,* 36:164-168 (June, 1952); Orville G. Manion, "An Application of Readability Formulas to Oral Communications," *Speech Monographs,* Abstracts of Theses, 21:151 (June, 1954); Kenneth A. Harwood, "Listenability and Readability," *Speech Monographs,* 22:49-53 (March 1955); Rudolph H. Vancura, "Flesch Readability Formula Applied to Television Programs," *The Journal of Applied Psychology,* 39:47-48 (February 1955); Beryl B. Blain, "Effects of Film Narration Type and of Listenability Level on Learning of Factual Information" (Unpublished Doctor's dissertation, Indiana University, 1956); Arnold G. Abrams, "The Relation of Listening and Reading Comprehension to Skill in Message Structuralization," *The Journal of Communication,* 16:116-125 (June 1966); Gerald M. Goldhaber and Carl H. Weaver, "Listener Comprehension of Compressed Speech When Difficulty, Rate of Presentation, and Sex of Listener are Varied," *Speech Monographs,* 35:20-25 (March 1968); and Robert Q. Young, "An Experimental Investigation of Reading and Listening Comprehension and of the Use of Readability Formulas as Measures of Listenability" (Unpublished Doctor's dissertation, Ohio University, 1972).

[6] See: Goldstein, *op. cit.;* Robert E. Jester and Robert M. W. Travers, "Comprehension of Connected Meaningful Discourse as a Function of Rate and Mode of Presentation," *The Journal of Educational Research,* 59:297-302 (March 1966); Goldhaber and Weaver, *op. cit.;* Emerson Foulke, "Listening Comprehension as a Function of Word Rate," *The Journal of Communication,* 18:198-206 (Sept. 1968); Charles M. Rossiter, "The Effects of Rate of Presentation on Listening Test Scores for Recall of Facts, Recall of Ideas, and Generation of Inferences" (Unpublished Doctor's dissertation, Ohio University, 1970); Ronald P. Carver, "Analysis of 'Chunked' Test Items as Measures of Reading and Listening Comprehension," *Journal of Educational Measurement,* 7:141-150 (Fall 1970); and Thomas G. Sticht, "Learning by Listening," *Language Comprehension and the Acquisition of Knowledge,* ed. John B. Carroll and Roy O. Freedle (Washington, D.C.: V.H. Winston & Sons, 1972).

[7] Edward Bliss, Jr., and John M. Patterson, *Writing News for Broadcast* (New York: Columbia University Press, 1971), p. 2.

[8] Klare, *The Measurement of Readability,* pp. 169-170.

ceive reproduced versions of typed news stories on a single sheet of paper and are allowed a minute and a half to read each story in their usual manner.

Paced Reading Group (PR)—Ss receive news stories one sentence at a time via a slide projector for a total time of a minute and half per story. This group is presented with each sentence for the same length of time as Ss in the oral condition.

Oral Group (OR)—Ss hear news stories via a tape recording for a total of a minute and a half per story. OR and PR groups are presented each sentence for the same length of time.

Control Group (CG)—Ss receive irrelevant stories via a normal written procedure when the other groups are receiving experimental presentations. This group receives the same pre- and posttests as the corresponding experimental group.

The purpose of the experiment was two-fold. First, it was predicted that a relationship between availability of repetition and increase in comprehension would be noted. Second, a difference in this effect in going from an "easy" readability to a "hard" readability level was predicted. Readability levels were determined by the use of the Dale-Chall formula.[9] The hard level in this study is the Dale-Chall formula's "College," or 13th to 15th grade level; the easy level is the 5th to 6th grade level.

Hypotheses. Research hypotheses are based, in effect, on the assumption that a paced reading condition, in which a reader is presented only one sentence at a time, via a projection screen, will more closely approximate the normal oral presentation situation than it does the normal reading situation in terms of comprehension. The hypotheses also take into account the "general principle" that "'difficult' material is more difficult to understand, and 'easy' materi-

[9] Edgar Dale and Jeanne S. Chall, "Formula for Predicting Readability," *Educational Research Bulletin*, 27:11-20, 28 (Jan. 21, 1948).

[10] James R. Young, *op. cit.*, p. 20.

al is easier to understand, when heard than when read."[10]

Specifically, the hypotheses are:

1. NR will show significantly less comprehension increase at the "easy" readability level than either PR or OR.

2. NR will show significantly greater comprehension increase at the "hard" readability level than either PR or OR.

3. All experimental groups (NR, PR, and OR) will show significantly greater comprehension increase than the corresponding CG.

Subjects. Subjects for this experiment were drawn from seven speech classes and two journalism classes at Marshall University during the first summer term in June of 1973. A total of 155 subjects took part in the pre-test and 140 were present for the experimental conditions and post-test. About 54% of the subjects were females and 46% males. About 74% were underclassmen and 26% were upperclassmen.

Subjects were randomly assigned to treatments. After the experiment was complete Ss were randomly discarded from groups so that all groups would be equal in size to the smallest group, in this case 15.

Messages and Measuring Instruments. Two news stories provided the content for the experimental messages in this study. Each story was rewritten twice—once to obtain a "hard" version and once to obtain an "easy" version. All four versions were the same length—264 words. The stories were about gun control in Great Britain and rising cotton prices in the United States. The hard version of the gun control story (GH) had a Dale-Chall raw score of 9.1746, placing it in the 13th-15th grade ("college") level. The hard version of the cotton price story (CH) had a Dale-Chall raw score of 9.1295, also placing it in the 13th-15th grade level. The easy version of the gun control story (GE) had a raw score of 5.1032, and was at the 5th-6th grade level. The easy version of the cotton price story (CE) had

a raw score of 5.0433, and also was at the 5th-6th grade level.[11]

These stories, then, comprised the experimental messages of the study. Also two "irrelevant" messages, both 264 words long, were produced for the control group. One of these messages dealt with Soviet uranium sales (UE) and was written at the 5th-6th grade level, raw score 5.7654; and the other was about current problems facing private colleges (PCH) and was at the 13th-15th grade level, raw score 9.060.

The oral messages were recorded using the researcher's voice at a speed of 176 wpm ±2 wpm. The reading time, then, was 1:30 ±2 seconds for each experimental message.

Each subject received a pre- and posttest over two of the experimental messages. For some this meant GE and CH; for others, GH and CE. The tests were produced by application of the cloze procedure to the experimental messages. An every-fifth-word deletion pattern was used; that is, the fifth word in each message was deleted and every fifth word thereafter until a total of 52 blanks was produced.[12]

Administration and Scoring. The different treatment groups were administered the stories in essentially the same fashion save for the modality difference. The NR group was told they had "about a minute and a half" to read a radio news story script and was then allowed to turn over a sheet of paper containing the entire passage. The PR group was told they would see a radio news story script flashed on the screen a sentence at a time and that the total story would take about a minute and a half to view. The OR presentation group was told they would listen to a radio news story via a tape recording and that it would take about a minute and a half. In all groups this procedure was the same for both the first story (either GE, or GH) and the second story (either CH, or CE).

Each subject generated four scores. Since each S was presented with two stories there were two pre-test measures

and two post-test measures for each individual.

Pre- and post-test scores for each individual were used to determine the gain score for each S for each story (G_I). These gain scores, defined as measures of comprehension increase, were used in the statistical analysis. A 4 (mode of presentation) X 2 (readability level) factorial analysis of variance was performed. Differences between individual pairs of means involved in the *a priori* hypotheses were then analyzed using *t*-tests.

Results and Conclusions

The overall analysis of variance in Table 1 shows that in the between-subject portion there is a significant difference between the four methods of presentation: CG, NR, PR, and OR. The combined means for the four CG measures was 1.98, for NR it was 8.05, for PR it was 7.08, and for OR it was 7.26.

The within-portion of the analysis indicates a significant difference between readability levels. The combined mean for both easy stories was 7.18, and for both hard stories 5.01.

No significant interaction between readability level and mode of presentation was found. The means and standard deviations for the eight modality x readability level combinations are reported in Table 2.

Differences between these eight means which were part of the *a priori* hypothesized relationships were analyzed via

[11] Readability analysis was done via the Denbow-Miller computer program which calculates the raw formula score for an inputed text. The initial version of this SNOBOL 4 program was written in 1970 by this author and Lawrence R. Miller of Ohio University. The program takes into account formula rules as set forth in: Edgar Dale and Jeanne S. Chall, "Formula for Predicting Readability: Instructions," *Educational Research Bulletin,* 27:27-54 (Feb. 18, 1948).

[12] All groups, including OR, were given pre- and posttests in the written mode. There is some indication that this might have resulted in lowered scores for the OR group since the "mode of expression" was not the same as the "mode of presentation." See: Milton W. Horowitz, "Organizational Processes Underlying Differences Between Listening and Reading as a Function of Complexity of Material," *The Journal of Communication,* 18:37-46 (March 1968).

Also it might be advisable for future studies to use all possible deletion-patterns of each story and then randomly assign them to appropriate Ss. This is necessary because some patterns alone may be easier or harder than the story as a whole and hence give a spurious index of comprehension.

TABLE 1

Analysis of Variance

Source	Sum of Squares	df	MS	F
Between Subjects				
Modality	1384.65	3	461.55	17.98*
Subjects Within Groups	2976.65	116	25.66	
Within Subjects				
Readability	283.84	1	283.84	16.81*
Modality X Readability	81.88	3	27.29	1.61
Readability X Subjects Within Groups[1]	1957.78	116	16.88	

*p<.0005

[1] Both a test of homogeneity of the covariance matrices of the independent groups, and a test of the assumption of homogeneous variance and homogeneous correlations between repeated measures were made. The former was significant ($X^2 = 19.86$, df=9, p $\lesssim$02) and the latter was not ($X^2 = 3.53$, df=1, p<.85). When there is evidence, as provided by the significant test above, which shows that the covariances between pairs of observations are not constant at all levels of the within subject factor Winer [p. 306 ff] suggests substituting 1 and p(n-1) for the usual degrees of freedom for the main within subject effects and substituting [(p-1) (q-1)] and [p(n-1) (q-1)] for the usual degrees of freedom for the treatment x within the interaction. In this case the only difference this makes is the readability level F where the degrees of freedom become 1 and 116, in place of 3 and 116. The F of 16.818 is, however, still significant at p<.0005.

TABLE 2

Means (and Standard Deviations) for Each Modality Within Each Readability Level

Readability Level	CG	NR	PR	OR
EASY	2.17 (4.55)	9.23 (4.17)	8.90 (5.04)	8.43 (4.49)
HARD	1.80 (2.76)	6.86 (5.68)	5.27 (4.26)	6.10 (4.76)

t-tests. The results of these comparisons show that within each readability level each modality evidences significantly greater comprehension increase than the corresponding control group, (p<.0005 in each case), but does not differ significantly from any of the other modalities. Since directional hypotheses were involved one-tailed tests were used. Because of these results Hypotheses 1 and 2 were rejected (null retained). However, hypothesis Number 3 was supported (null rejected).

It would appear, then, that readability ($\Sigma G_I/N$ for each NR and PR group) is equivalent to listenability ($\Sigma G_I/N$ for each corresponding OR group). That is, the major dependent variable of comprehension increase does not differ significantly across the independent variable of modality within the independent variable of readability level.

Relationship of listenability to the Dale-Chall readability formula score which determined assignment to easy and hard categories can be looked at in the following manner. Since the combined easy stories led to significantly greater comprehension gain than the combined hard stories, and since there was no significant interaction between modality and readability level, it appears that over both contents formula grade-level placement predicted accurately the relative difficulty of obtaining increases in both listening (OR) and reading (NR and PR) comprehension.

Also the lack of a significant difference between PR, NR and OR means within each readability level seems to indicate that with the kind of material used in this study ability to re-read earlier sentences did not increase reading gain. That is, the stated assumption underlying hypotheses one and two that repetition accounts to a significant extent for differences between listenability and readability is not supported by the results.

A recent study by R.Q. Young obtained similar results,[13] especially in terms of OR and PR comparison at various readability levels. Both studies used relatively short messages written for listening. It is possible that these materials eliminated the kinds of construction which may have contributed to the differences found in some of the earlier studies. Further, it is conceivable that longer passages which would put a greater burden on memory span would also increase the importance of a repetition factor.

These studies do show, however, that with short news stories in modern style[14] repetition is as important (or unimportant) to the broadcast journalist as to his newspaper colleague.[15] To the extent these results are valid, the broadcast journalist does not need to be *overly* concerned with the adage of the trade: "Tell 'em what you're going to tell 'em. Tell 'em. And then tell 'em what you've told 'em."

[13] Robert Q. Young, *op. cit.,* It should be pointed out that Young used a multiple-choice comprehension test in an after-only design. His dependent variable then was comprehension, not comprehension gain.

[14] Modern newspaper style has, for the most part, adopted many of the tenets of broadcast style: shorter sentences, less anaphora and a more conversational tone.

[15] Specifically the results could be summarized in the following way. (1) When material written "for the ear" is used and time of presentation controlled for, comprehension increase does not differ across modalities. This relationship tends to hold true *within* readability levels as gauged by the Dale-Chall formula. (2) The Dale-Chall readability formula is as adequate in measuring listenability as it is in measuring readability. (3) More generally, the results indicate that there is no inherent limitation of comprehension due to the modality used by the broadcast journalist when relatively short material is delivered at a speed similar to that used in this experiment. Thus the underlying assumption in regard to the value of repetition is not supported in the present context.