

# **The ReSULT: The Revised Shortened Utley Sentence Lipreading Test**

Claudia D. Updike  
*Ball State University*

The purpose of this study was to reduce the length of the Utley Sentence Lipreading Test (ULT) without affecting its reliability and validity. Twenty sentences from the ULT were selected via item analysis as being most highly correlated with the total score. The revised test (ReSULT) and the ULT were administered to 213 normal-hearing young adults. Significant correlations were obtained between the ReSULT and the ULT, and between the two forms of the ReSULT. High internal consistency and test-retest reliability were established. It was concluded that the ReSULT is a reliable alternative to the Utley Sentence Lipreading Test.

One concern of speech-language pathologists and rehabilitative audiologists involved in speechreading instruction is the determination of speechreading performance of the hearing-impaired individual. Tests should be reliable and valid. If a speechreading test can be administered in a short length of time, yet still retain its reliability and validity, the margin of error due to client fatigue may be reduced, and the administrator's additional time may be utilized for other purposes. Such a test would be thus less taxing for the client and more cost effective.

Numerous tests have been devised to evaluate speechreading ability, ranging from nonsense syllables to short paragraphs (e.g., Binnie, Jackson & Montgomery, 1976; Donnelly & Marshall, 1967; Johnson, 1976; Nielsen, 1970). Of the available tests, the best known and most widely used is the Utley Lipreading Test (Jeffers & Barley, 1981). The reliability of the Utley Test has been evaluated repeatedly using large sample populations; a coefficient of .94 for the Utley Sentence Test Form A, with Form B has been obtained (DiCarlo & Kataja, 1951; Utley, 1946).

Both forms of the Utley Test Part II, the Sentence test, contain 31 sentences thought to be a "representative sample of oral language commonly used by individuals in all parts of the country" (Utley, 1946, p. 110). From personal

---

Claudia D. Updike, PhD, is an Associate Professor of Audiology in the Department of Speech Pathology & Audiology at Ball State University, Muncie, IN 47306.

experience, it has been learned that most hearing-impaired individuals, particularly the elderly, have difficulty maintaining optimum attention throughout Part II of the test.

Research has indicated that a large number of items in the Utley Test may be nonfunctional. Heider (1947) administered the entire Utley Lipreading Test to Middle School (ages 9:9 to 14:3) and Upper School (ages 13:2 to 19:1) children of Clarke School and she concluded that half of the items should be discarded because they were so difficult that they did not discriminate between levels of speechreading ability. DiCarlo and Kataja (1951) administered the Utley Test to a group of normal-hearing adults to determine whether or not the Utley Test is a valid and reliable instrument to test speechreading. Despite the high correlation of .94 between forms, the investigators observed that 50% of the items elicited responses from fewer than 3% of the subjects. Furthermore, they found that the highest score attained was only 39% of the maximum, while the mean score was only 19% of the maximum.

DiCarlo and Kataja concluded that the Utley Test consists of a considerable number of undiscriminating items which contribute nothing to the test. They also concluded that the Utley Test discourages even the best speechreaders and it seems to "test the ability to tolerate frustration and persistent failure" (DiCarlo & Kataja, 1951, p. 240). They speculated that the Utley Test might be shortened without negative effects. Despite the recommendation of the authors of that study to devise a more efficient version of the Utley Lipreading Test, there seems to have been no attempt in the past 35 years to carry out this much needed work. Therefore, the purpose of this study was to evaluate the reliability of a shortened version of Part II of the Utley Test as a measure of speechreading performance.

## PART 1

### Method

*Subjects.* A group of 50 normal-hearing college students, ages 18 to 25 years, served as subjects in Part 1 of this study. All subjects passed a vision screening test (Optec 1000 DMV Vision Test). The criterion for passing was 20/40 because research indicates that visual acuity of at least 20/80 is necessary for success in speechreading (Romano & Berlow, 1974). The subjects also passed a hearing screening test at 25 dB HL at the octave frequencies from 250 to 8000 Hz.

*Test material and administration.* Videotapes were produced of a woman, with a General American dialect, speaking the Utley Test, Part II Sentences, Forms A and B. The image recorded was a front view of the talker, including her head and shoulders. High-intensity studio lighting provided a clear view of the talker's face and mouth. A Panasonic WV-3890B portable color camera was used and the videotapes were presented using a Panasonic AG 6300 videotape player with an RCA XL-100 Solid State monitor (screen size: 21 in. by 16 in.).

Both Forms A and B of the Utley Sentence Lipreading Test were administered. To control for order effect, Form A was presented first to one half of the group,

and Form B was presented first to the other half. Test administration was conducted in a room with typical classroom lighting in groups of up to seven persons. The distance from talker/monitor to speechreaders did not exceed 8 ft. because research indicates that speechreading scores diminish about .8% per foot with increasing distance beyond six feet (Erber, 1971). Each test item was presented once without sound, and the subjects wrote down their responses. They were instructed to write as much of the sentences/phrases that they recognized. One point was given per correct word. If the subject used contractions (e.g., "I've" for "I have"), two points were awarded for these structures. Correct word order was required.

### Results and Discussion

Mean raw scores were 82.3 on Form A and 78.3 on Form B (Table 1). A Pearson product-moment correlation coefficient of .81 (Table 2) was obtained between Form A and Form B. While significant at the .05 level, the correlation between the two forms of the Utley Sentence Test was too low for the two forms to be considered equivalent alternative forms. Because a shorter test derived from either of the original two forms of the Utley Sentence Lipreading Test would have had even lower between-list correlations (Nunnally, 1978), Forms A and B were combined.

An item analysis was performed on the data from Forms A and B combined. Table 3 gives the correlation of each item to the overall combined test. The internal consistency (Cronbach's alpha, Carmines & Zeller, 1979) of the Utley Test (Forms A and B combined) was high (.92%). Several test items had very low correlations to the overall test. These results agree with the data obtained by DiCarlo and Kataja (1951).

The 20 items which had the highest correlation coefficients were rank-ordered

**Table 1**  
Number of Words Correct on the Utley Sentence Test (Utley)  
and the Revised Shortened Utley Sentence Lipreading Test (ReSULT)  
by Normal-Hearing Young Adults

Subjects		Utley		ReSULT	
		Form A	Form B	Form A	Form B
Part 1 ( <i>N</i> = 50)	<i>M</i>	82.3	78.3		
	<i>SD</i>	(19.0)	(15.8)		
Part 2 ( <i>N</i> = 213)	<i>M</i>	79.2	76.0	26.0	28.5
	<i>SD</i>	(19.2)	(19.9)	(9.7)	(10.4)
Part 3 ( <i>N</i> = 36)	Initial Test			24.6	26.3
		<i>M</i>		(9.8)	(7.8)
Retest	<i>M</i>			26.7	28.9
	<i>SD</i>			(9.1)	(7.3)

**Table 2**

Pearson Product-Moment Correlation Coefficients Between Scores on the Utley Sentence Test (Utley) and the Revised Shortened Utley Sentence Lipreading Test (ReSULT) for Normal-Hearing Young Adults

Variable 1 (Test/Form)	Variable 2 (Test/Form)	Subjects		
		Part 1 (N = 50)	Part 2 (N = 213)	Part 3 (N = 36)
Utley-A	Utley-B	.81	.82	
ReSULT-A	Utley A & B		.89	
ReSULT-B	Utley A & B		.91	
ReSULT-A	ReSULT-B		.82	
ReSULT-A (Initial test)	ReSULT-A Retest			.90
ReSULT-B Initial test	ReSULT-B Retest			.86

**Table 3**

Reliability Analysis for the Utley Sentence Lipreading Test, Forms A and B Combined

Item Number	Correlation of Item with Total Test	Alpha if Item Deleted
<b>Form A</b>		
1	.3274	.9237
2	.2699	.9238
3	.4726	.9229
4	.3058	.9236
5	.4573	.9227
6	.3841	.9227
7	.2954	.9233
8	.5352	.9237
9	.3390	.9218
10	.4314	.9228
11	.3727	.9232
12	.4184	.9229
13	.6257	.9225
14	.1622	.9243
15	.4802	.9230
16	.5763	.9214
17	.2649	.9239
18	.1904	.9242
19	.5028	.9223
20	.5600	.9216
21	.6217	.9209
22	.3844	.9233

*Continued on next page*

Table 3 Continued

Item Number	Correlation of Item with Total Test	Alpha if Item Deleted
23	.5253	.9227
24	.3271	.9235
25	.4612	.9228
26	.6083	.9213
27	.5468	.9218
28	.3470	.9234
29	.5203	.9220
30	.5316	.9219
31	.3081	.9236
<b>Form B</b>		
1	.0672	.9244
2	.3663	.9233
3	.3291	.9235
4	.2390	.9240
5	.1445	.9243
6	.3026	.9240
7	.3334	.9235
8	.3107	.9237
9	.5005	.9225
10	.3478	.9235
11	.2815	.9238
12	.4118	.9229
13	.4446	.9228
14	.3784	.9233
15	.3481	.9234
16	.1529	.9243
17	.3226	.9235
18	.3021	.9236
19	.5110	.9221
20	.3625	.9234
21	.5243	.9226
22	.5175	.9221
23	.5219	.9226
24	.4576	.9226
25	.4368	.9232
26	.4182	.9229
27	.4276	.9228
28	.4749	.9230
29	.4993	.9222
30	.2341	.9244
31	.4608	.9235

Note. Alpha = .92 for Forms A and B combined.

and divided into two lists. The item with the highest correlation was placed in List A, the second highest in List B, the third highest in List B, the fourth highest in List A, and so on. List A was then named Revised Shortened Utley

Sentence Lipreading Test, Form A (ReSULT-A). List B was named Revised Shortened Utley Sentence Lipreading Test, Form B (ReSULT-B) (Appendix A).

## PART 2

Scores from a second group of subjects were collected to determine whether the items chosen to make up the revised test were significantly correlated with the combined lists of the original test. The purpose was, therefore, to evaluate the criterion validity and the internal consistency of the ReSULT.

### Method

*Subjects.* Subjects were 213 normal-hearing college students, ages 18 to 25 years. Each subject passed the vision and hearing screening tests described in Part 1. None of the subjects in Part 2 had participated in Part 1.

*Test material and administration.* The same equipment and materials utilized in Part 1 were employed in Part 2 of the study. In addition, videotapes were produced of a woman speaking the sentences in ReSULT-A and ReSULT-B. This was the same talker as for Part 1. It is recognized that the new recordings of the test stimuli for the ReSULT may have represented slightly different tokens of the stimulus items than appeared on the first recordings of the Utley Sentence Lipreading Test because few talkers can precisely reproduce visible articulation.

The same procedure was used in Part 2 as in Part 1. In addition, the ReSULT-A and ReSULT-B were administered to each subject. To control for order effects, one half of the group received the ReSULT-A followed by Utley Form A first, then ReSULT-B followed by Utley Form B; the other half of the group received ReSULT-B, followed by Utley Form B first, then ReSULT-A followed by Utley Form A. The same test conditions and scoring method were used in Part 2 as in Part 1.

### Results and Discussion

Mean raw scores were 79.2 and 76.0 on Form A and Form B of the Utley Sentence Test Part II, respectively (Table 1). These scores were slightly lower than those obtained in Part 1. There was a .82 correlation between scores on Form A and those on Form B. This correlation was in good agreement with the results obtained with the young normal-hearing adults in Part 1 of the study.

Mean raw scores were 26.0 on the ReSULT-A and 28.5 on the ReSULT-B (Table 1). Pearson product-moment correlation coefficients for scores on the Utley Forms A and B combined were  $r = .89$  with the ReSULT-A and  $r = .91$  with the ReSULT-B ( $p < .01$ ) (Table 2). Cronbach's alpha for the ReSULT-A was .82 and for the ReSULT-B it was .85. Table 4 presents the item total correlation and internal validity information for the ReSULT Forms A and B. As a measure of the between-list reliability of the two revised forms, a Pearson product-moment correlation coefficient (Table 2) of .82 was calculated ( $p < .01$ ) (Table 1).

Raw scores, Z-scores, and percentile ranks are presented in Appendix B. These

data can be used to compare the performance of an individual on the ReSULT Forms A and B with that of a large group of normal-hearing young adults. Additional norms need to be developed for the hearing-impaired population for clinical use.

### PART 3

Test-retest reliability data are critical to ensure that any change in a speech-reading score following rehabilitation is due to treatment rather than to the lack of reliability inherent within the test. Therefore, Part 3 of the study was conducted to evaluate the test-retest reliability of the ReSULT Forms A and B.

#### Method

*Subjects.* Thirty-six normal-hearing college students, ages 18 to 25 years, participated. All subjects passed the vision and hearing screening tests described in Part 1. None of these subjects had participated in either of the first two parts

**Table 4**  
Reliability Analysis for the ReSULT Forms A and B

Item Number	Correlation of Item with Total Test	Alpha if Item Deleted
<b>ReSULT-A*</b>		
1	.36	.82
2	.33	.82
3	.44	.81
4	.67	.78
5	.65	.78
6	.73	.77
7	.58	.80
8	.48	.80
9	.27	.82
10	.51	.80
<b>ReSULT-B<sup>b</sup></b>		
1	.54	.82
2	.68	.80
3	.59	.81
4	.53	.82
5	.43	.83
6	.49	.83
7	.42	.84
8	.54	.82
9	.36	.84
10	.54	.82

\*Alpha = .82. <sup>b</sup>Alpha = .85.

of the study.

*Test materials and administration.* The ReSULT Forms A and B were administered using equipment identical to that used in Parts 1 and 2. Form A was presented first to one half of the group, while Form B was presented first to the remaining half. One week later, the ReSULT was readministered. Each subject received the retests in the same order as the initial tests. The test items were not discussed with any of the subjects until after the retest condition. Test conditions and scoring were similar to those previously described.

### **Results and Discussion**

Mean raw scores obtained on the two administrations of the ReSULT Forms A and B are provided in Table 1. On the retest, the mean scores were nominally higher, but  $t$  tests confirmed that there were no significant differences between test and retest scores ( $t = -1.19$ ,  $df = 71$  for ReSULT-A;  $t = -1.47$ ,  $df = 71$ , for ReSULT-B;  $p > .05$ ). The scores were similar to those obtained for the 213 subjects who participated in Part 2. Test-retest reliability was computed using Pearson product-moment correlations:  $r = .90$  for ReSULT-A, and  $r = .86$  for ReSULT-B. These values are considered to be acceptable for test-retest reliability (Nunnally, 1978).

## **DISCUSSION AND CONCLUSIONS**

The purpose of this study was to develop a shortened version of the Utley Sentence Test, without adversely affecting its reliability. The results of Part 1 clearly indicate that the Utley Sentence Lipreading Test contains numerous un-discriminating items regarding speechreading performance. This suggests that the Utley Sentence Lipreading Test could be shortened without affecting its reliability. The two forms of the shorter ReSULT were highly correlated and their internal consistency was moderately high. These data suggest that the ReSULT is a reliable, time-efficient test.

It is acknowledged that the data collected in this study were from young, normal-hearing adults. In order for the ReSULT to be of clinical use with special populations, such as the elderly hearing-impaired population, more normative data are needed.

The statistically high reliabilities of the ReSULT-A and ReSULT-B suggest that this test may be useful in evaluating improvement in speechreading performance resulting from speechreading instruction. Either form of the ReSULT may be used as a pre- or a posttest.

The ReSULT has much potential for clinical use. The provision of normative data for this test allows the clinician to compare the speechreading skills of a particular student with normal-hearing performance, a practice used by Utley to generate the original Utley Sentence Lipreading Test. Furthermore, the ReSULT offers the ability to evaluate change in speechreading ability by having two forms.

The ReSULT also has potential as a research tool. It can be used as a time-

efficient method to evaluate the effectiveness of various techniques of speechreading instruction. It can also be utilized to evaluate the effects of internal or external factors on speechreading performance. In summary, the ReSULT offers a reliable alternative to the Utley Lipreading Test, and has potential value as a clinical and research tool.

### REFERENCES

- Binnie, C., Jackson, P., & Montgomery, A. (1976). Visual intelligibility of consonants: A lipreading screening test with implications for aural rehabilitation. *Journal of Speech and Hearing Disorders*, 14, 530-539.
- Carmines, E., & Zeller, R. (1979). *Reliability and validity assessment*. Beverly Hills, CA: Sage Publications.
- DiCarlo, L.M., & Kataja, R. (1951). An analysis of the Utley Lipreading Test. *Journal of Speech and Hearing Disorders*, 16, 226-240.
- Donnelly, K., & Marshall, W. (1967). Development of a multiple choice test of lipreading. *Journal of Speech and Hearing Research*, 10, 565-569.
- Erber, N.P. (1971). Effects of distance on the visual reception of speech. *Journal of Speech and Hearing Research*, 14, 848-857.
- Heider, G. (1947). The Utley Lipreading Test. *Volta Review*, 49, 457-458, 488-490.
- Jeffers, J., & Barley, M. (1981). *Speechreading (Lipreading)*. Springfield, IL: Charles Thomas.
- Johnson, D. (1976). Communication characteristics of a young deaf adult population: Techniques for evaluating their communicative skills. *American Annals of the Deaf*, 121, 409-424.
- Nielsen, H. (1970). Measurement of visual speech comprehension. *Journal of Speech and Hearing Research*, 13, 856-860.
- Nunnally, J.C. (1978). *Psychometric theory* (2nd ed.). New York: McGraw Hill.
- Romano, P.E., & Berlow, S. (1974). Visual requirements for lip reading. *American Annals of the Deaf*, 119, 383-386.
- Utley, J. (1946). A Test of lipreading ability. *Journal of Speech Disorders*, 11, 109-116.

### APPENDIX A

#### REVISED SHORTENED UTLEY SENTENCE LIPREADING TEST (ReSULT-A)

1. I have forgotten.
2. Good night.
3. Did you have a good time?
4. She was home last week.
5. Of course.
6. We got home at six o'clock.
7. We drove to the country.
8. Did you like it?
9. It is around four o'clock.
10. The thermometer says twenty above.

Total possible score = 43.

#### REVISED SHORTENED UTLEY SENTENCE LIPREADING TEST (ReSULT-B)

1. Keep your eye on the ball.
2. The train runs every hour.
3. You had better go slow.

4. I don't know if I can.
5. What does the paper say about the weather?
6. Do you understand?
7. I had rather go now.
8. I cannot stand him.
9. Excuse me.
10. Come again.

Total possible score = 46.

### APPENDIX B

#### SCORE TRANSFORMATIONS FROM THE RESULTS OF ADMINISTERING THE REVISED SHORTENED UTLEY SENTENCE LIPREADING TEST (ReSULT) TO 213 NORMAL-HEARING YOUNG ADULTS

Raw Score	Form A		Form B	
	Z-Score	Percentile Rank	Z-Score	Percentile Rank
46			1.68	95
45			1.58	94
44			1.49	93
43	1.76	97	1.39	92
42	1.66	95	1.29	91
41	1.56	94	1.20	89
40	1.45	93	1.10	87
39	1.35	92	1.01	86
38	1.24	89	.91	84
37	1.14	87	.81	79
36	1.04	86	.72	77
35	.93	84	.62	74
34	.83	79	.53	72
33	.73	77	.43	66
32	.62	74	.33	62
31	.52	70	.24	58
30	.42	66	.14	56
29	.31	64	.05	51
28	.21	57	-.05	49
27	.10	52	-.15	47
26	.00	50	-.24	43
25	-.10	46	-.34	34
24	-.21	43	-.43	32
23	-.31	38	-.53	30
22	-.41	34	-.63	26
21	-.52	30	-.72	24
20	-.62	28	-.82	21
19	-.72	25	-.91	16
18	-.83	19	-1.01	14
17	-.93	16	-1.10	13

*Continued on next page*

Appendix B *Continued*

Raw Score	Form A		Form B	
	Z-Score	Percentile Rank	Z-Score	Percentile Rank
16	-1.04	14	-1.20	11
15	-1.14	12	-1.30	9
14	-1.24	11	-1.39	8
13	-1.35	9	-1.49	6
12	-1.45	8	-1.59	5
11	-1.55	7	-1.68	4
10	-1.66	6	-1.78	3
9	-1.76	5	-1.87	2
8	-1.86	4	-1.97	2
7	-1.97	3	-2.07	2
6	-2.07	2	-2.16	1
5	-2.18	1	-2.26	1
4	-2.28	1	-2.35	1
3	-2.38	-1	-2.45	-1
2	-2.49	-1	-2.55	-1
1	-2.59	-1	-2.64	-1
0	-2.69	-1	-2.74	-1