

AN INQUIRY INTO POSSIBLE APPLICATION OF "CLOZE-PROCEDURE" AS A DIAGNOSTIC AND THERAPEUTIC TECHNIQUE IN CASES OF ADULT DYSPHASIA

> Thesis for degree of Master of Arts Michigan State University

> > WILLIAM HOWARD HAAS

1112 - 5

LIBRARY Michigan State University

AN INQUIRY INTO POSSIBLE APPLICATION OF "CLOZE-PROCEDURE" AS A DIAGNOSTIC AND THERAPEUTIC TECHNIQUE IN CASES OF ADULT DYSPHASIA

by

WILLIAM HOWARD HAAS

## A THESIS

Submitted to the College of Communication Arts Michigan State University of Agriculture and Applied Science in partial Fulfillment of the requirements for the degree of

MASTER OF ARTS

Department of Speech

1960 APPROVED:

## ABSTRACT

(William H. Haas)

"Cloze-Procedure" originated and validated by Wilson L. Taylor at the University of Illinois, has been reported to indicate working intelligence, attention span, knowledge and learning of a subject, perception, language skill, listening comprehension, and reading comprehension.

The purpose of this investigation was to explore possible application of this concept to cases of adult dysphasia. The specific area of measuring learning was chosen for experimentation. The following hypothesis was stated: Comparison of the results of pre-indoctrination and post-indoctrination scores obtained from a given text will indicate post scores are higher than pre scores; therefore, measuring the degree of learning that must have taken place.

The collected data consisted of the responses of 15 subjects to 84 word deletions on five reading tasks. Experimentation was initiated through three trials at "closure." Trial One was instigated to measure existing knowledge for the topics in the investigation. Trial Two was accomplished to test the ability of this technique to quantitatively measure a degree of learning that took

## (William H. Haas)

place following indoctrination sessions. A third attempt was provided in order to ascertain the instrument's sensitivity to measure recall after the delay of a one week time factor.

An analysis of the results indicated a substantial increase in the raw scores and mean gain between Trials One and Two. However, a comparison between the scores of Trials Two and Three, was not as obvious; therefore, a statistical procedure was applied in the form of a one tail simple  $\underline{t}$  test. The results were highly significant beyond the .01 level of confidence. It was concluded that "cloze-procedure" was successful in measuring learning, within the conditions of the study, and that the technique was sensitive to the degree of measuring delayed reaction to recall.

The multiplicity of applications this technique seemingly possesses was suggested. The concept of "closure" was discussed as a therapeutic aid, an evaluation instrument, and a potential resource for future research.

Charles Pedrey, Ph.D. Major Professor

## TABLE OF CONTENTS

I.	ACKNOWLEDGMENTS	iv
II.	LIST OF TABLES	v
III.	INTRODUCTION	l
IV.	PURPOSES OF STUDY	4
V.	THE CLOZE PROCEDURE	7
VI.	HYPOTHESIS AND EXPERIMENTAL DESIGN	12
VII.	RESULTS	19
VIII.	CONCLUSIONS AND DISCUSSION	24
IX.	SUMMARY	29
X.	APPENDIX	31
XI.	BIBLIOGRAPHY	42

## ACKNOWLEDGMENTS

Foremost, I would like to express my sincere appreciation to Dr. Charles Pedrey, Director of the Speech Clinic at Michigan State University, for his encouragement, guidance, and kindness in planning and developing this paper.

Special thanks are extended to Dr. Max Nelson, Director of the Hearing Clinic at Michigan State University, and Mrs. David Hayes, Psychologist in Ft. Worth, Texas, for their critical comments and suggestions.

A special acknowledgment is extended to Mr. William J. Haywood, Consultant in Clinical Psychology at the Society for Crippled Children and Adults' Rehabilitation Center in Ft. Worth, Texas, for his invaluable contribution to the analysis of the statistical data for this study.

iv

# LIST OF TABLES

Table		Page
1.	Order of Subject Participation in Individual Testing	18
2.	Mean Number of Correct Responses on Three Attempts at Closure	20
3.	Mean Difference, Standard Deviations, and Standard Error Revealed Through Three Trials at Closure	22
4.	Standard Error and <u>t</u> Score Between Trials Two and Three	23
5.	Raw Scores of 15 Subjects in Three Attempts at "closure"	31

лалтала 1 алтала 1

•••••

•

### INTRODUCTION

Hospitals and rehabilitation centers throughout the United States report that the referral incidence of adults with aphasia is constantly increasing. This higher incidence of referral is due, partially, (1) to more exacting medical diagnosis of aphasia, (2) increases in strokes and accidents caused by the pressure of fastpaced living, and (3) progress of medical science in decreasing the death rate immediately following the cortical insult.<sup>1</sup> With this subsequent influx of aphasic patients into speech centers throughout the country, it becomes concurrently imperative that the speech therapist re-evaluate, closely define, and improve approaches to be used in therapy.

Information regarding therapy techniques for the initial stages of building a basic language function is much more readily found than information for measuring and sophisticating a language function that already exists. Many speech therapists working in the "vocationally oriented" rehabilitation center often find themselves in need of this type of information. For our discussion here, therefore, we might classify aphasic patients into

<sup>1</sup>Nancy E. Wood, "Helping the Aphasic Adult," Journal of Rehabilitation, 22:7, (January, 1956). p. 7.

two categories: (1) those with little or no basic language function, and (2) those with a basic language function in need of sophistication. It is with the latter that we will have primary interest.

Linguists and psycho-linguists have urged our profession to utilize more of the linguistic concepts when dealing with this type of language behavior.<sup>2</sup> They tell us that they have a wealth of information which seemingly would have great potential for application to aphasia problems.

A survey of the literature brings forth the gestalt notion of "clozure." In fact, a method to measure this in regard to language behavior has been introduced and validated by William L. Taylor and associates. It is called the "cloze-procedure."<sup>3</sup>

The term "cloze" is derived from the idea of closure, or the tendency to fill in a missing link in a well-structured whole. Given the sequence, "dogs bark and \_\_\_\_\_\_ meow," most readers would immediately supply the missing "cats." As the actual procedure has been worked out for application, the experimenter deletes every nth word in a text, leaving the equal sized blanks in their places and the subjects read through the passage filling in

<sup>2</sup>M. H. Sargill, "Modern Linguistics: Recovery from Aphasia," Journal of Speech and Hearing Disorders, 19 (December, 1954), p. 507.

<sup>3</sup>Wilson L. Taylor, "Application of 'Cloze' and Entropy Measures to the Study of Contextual Constraint in Sample of Continuous Prose," (unpublished Ph.D. dissertation, University of Illinois, 1954).

the missing words. The more closely the totality of sequential cues in the passage elicits at each test point the same word thought as intended, the higher will be the subject's "cloze" score, or communication success.<sup>4</sup>

In the above sequence, the reader could not have guessed "cats" unless he: (1) had functional native ability, (2) attended to the message, (3) was familiar with the words and grammar of the message, (4) associated the message with feline sounds, (5) understood the whole meaning of the mutilated sentence, and (6) had language habits which made him use the same word the message called for. "The 'guess' depended on his intelligence, attention, knowledge, perception, language habits, and comprehension."<sup>5</sup>

The significant advantage of Taylor's "clozeprocedure" is that it taps simultaneously all of the complex determinants affecting word choice, both at the various levels of organization and through stretches in context. "Cloze" scores seem to depend on most, and perhaps all, of the many factors which have significant influence on communication; receptive, expressive, and intellectual abilities.

<sup>5</sup>Wilson L. Taylor, "Cloze Procedure," <u>Agrisearch</u>, 2:2 (February, 1956), p. 3.

<sup>&</sup>lt;sup>4</sup><u>Ibid</u>., p. 68.

## PURPOSES OF STUDY

As pointed out above, the procedure has been reported to indicate working intelligence, attention span, knowledge and learning of a subject, perception, language skill, listening comprehension, and reading comprehension.

If "cloze-procedure" is successful in measuring these areas with normal adults, what application can be made with cases of adult dysphasia? Obviously, in a paper of this scope, it is impossible to investigate all areas with integrity. Consequently, this writer has chosen to study the "procedure's" usefulness in measuring the amount of learning exhibited by a given patient or group of patients for a particular text.

<u>HYPOTHESIS</u>. Comparison of the results of pre and post "cloze" scores obtained from a given text will indicate post-indoctrination scores are higher than preindoctrination scores; therefore, measuring the amount of learning that must have taken place in cases of adult aphasia.

A secondary purpose is to discuss other possible application of the concept, specifically: its potential as a therapy method, an evaluation instrument, and its implication for future research.

To facilitate the discussion of these objectives this investigation is arranged into the following areas:

(1) definitions and delimitations, (2) historical discussion of the methodology, research, and previous application of "cloze-procedure," (3) discussion of experimental design, hypothesis, and results, (4) other potential uses of the technique, and (5) conclusion.

Definitions and delimitations. There is great variance in the use of the term aphasia among the professionals in the field, and by the laymen. There is agreement that the term should be limited to language disorders which derive from organic impairments. However. a language disturbance caused by deafness, blindness, or by various psychological disorders should not be referred to as aphasia. Dysarthria, although neurological in nature. is not included because faulty symbolizing is not characteristic.<sup>6</sup> Therefore, for the needs of this writing, aphasia is defined as the loss of symbolic language as the result of brain damage. For the purposes here, there is no need for classification definitions.

<u>Dysphasia</u> is a term which will be used to designate a milder degree of such disturbance. Under this category will be the patient who exhibits a useful language function.

<sup>6</sup>Helmer R. Myklebust, "Aphasia in Children: Language Development and Language Pathology," <u>Handbook of Speech</u> <u>Pathology</u>, ed. Lee E. Travis (New York: Appleton-Century-Crofts, Inc. 1957) p. 507.

<u>A basic language function</u> will be said to have been fulfilled when the patient: (1) obtains a vocabulary of approximately 100 words which he can both understand and verbally express, (2) uses short phrases or sentences to express his needs, (3) can read sentences and short paragraphs with comprehension, and (4) can express himself in writing using at least single words. Absolute accuracy, of course, should not be a factor in all cases.

"Cloze-procedure" is a technique that provides a measure of the degree of correspondence between the language habits used by a human transmitter while expressing a message and those employed by a human receiver while interpreting this message. It does so by intercepting and mutilating a message between transmitter and receiver, and asking the receiver to make it whole again.<sup>7</sup> A <u>cloze score</u> is simply the number of times deleted elements of a text are successfully anticipated and replaced in context.<sup>8</sup>

7Taylor, "Application of 'Cloze' and Entropy Measures...," p. 68. <u>8Ibid</u>.

### THE CLOZE PROCEDURE

Historical review. The "Cloze-procedure" was first instroduced in 1953 at the University of Illinois.<sup>9</sup> Initially, it was regarded mainly as an improved way of measuring the readability, or ease of reading by an individual or group of individuals, of given texts covering specific subject matter. The great bulk of studies were devoted to the procedure's readability uses.<sup>10</sup>

Taylor demonstrated the feasability of his technique for this purpose. In fact, it "produced" much more satisfactorily than either the Flesch or Dale-Chall formulas, traditionally used as demonstrations by the authors of these formulas in the same way, but on specialized texts it alone yielded sensible results. For example, both Flesch and Dale-Chall indicate a passage from Gertrude Stein as being easy. Taylor's procedure shows Stein, more appropriately, as <u>not</u> easyi In other words, this technique takes into account the unpredictable expressive style characteristic of Stein.<sup>11</sup>

<sup>&</sup>lt;sup>9</sup>Wilson L. Taylor, "Cloze-Procedure: A New Tool for Measuring Readability," <u>Journalism Quartly</u>, XXX (Fall, 1953) 415-533.

<sup>10&</sup>lt;sub>Ibid</sub>.

<sup>11&</sup>lt;sub>Ibid</sub>. p. 425.

As a measure of degree of comprehension, the "procedure" also has enjoyed success. In a very carefully designed experiment using Air Force personnel for which comprehension tests were already available, it was found that "cloze" scores correlated very highly with initial comprehension scores (pre-indoctrination) and also predicted final comprehension (post-indoctrination).

Approximately 150 Sampson Air Base trainees were used for this study. The experiment was administered in two parts separated by a week's time. As a first step, the subjects were given the "cloze" form, then a carefully constructed "before" comprehension test. This was done to measure their existing knowledge of the article's contents. Secondly, the subjects were given the article to study. Then, an "after" comprehension test was given. And lastly, another "cloze" form was administered. The results showed the "cloze" and traditional score differences were statistically significant to above the .001 level of confidence. Further, scores using "closure" were positively correlated with general intelligence at .74. Intelligence quotients were obtained from the Armed Forces Qualification Test.<sup>12</sup>

The usefulness of the technique has not been limited to the English language. The procedure was found

<sup>&</sup>lt;sup>12</sup>Wilson L. Taylor, "The Cloze Procedure: How it predicts Comprehension and Intelligence of Military Personnel," <u>Technical Memorandum No. 13</u>, Urbana, Illinois: Division of Communications, University of Illinois, (1953).

to be successful in Korean, a language written in a mixture of phonetic symbols and Chinese characters.<sup>13</sup>

Other studies have shown the "cloze-procedure's" effectiveness in measuring the degree of listening 14 comprehension, as well as reading comprehension.

Kerrick, at the University of California, has recently directed research which has shown that there is a significant correlation between how well readers say they like an article and their "cloze" scores on them.<sup>15</sup>

Taylor reports that studies have been undertaken which use still broader application of the "procedure" in other specialized areas. For example, individual differences and aptitudes have been studied in relation to reading comprehension.<sup>16</sup>

To this writer's knowledge, there has been no application of this technique to the language behavior of the dysphasic patient.

Methodology of administration and scoring. Methodology for application of the "procedure" has been formulated by Taylor through a series of pilot studies.

<sup>16</sup>Taylor, "Cloze Readability Scores as Indices of Individual Differences in Comprehension and Aptitude," Journal of Applied Psychology, XLI (February, 1957), 19-26.

<sup>17</sup>Taylor, "Cloze-Procedure: A New Tool...," p. 415.

<sup>&</sup>lt;sup>13</sup>Wilson L. Taylor, "Recent Developments in the Use of 'Cloze-Procedure'." <u>Journalism Quarterly</u>, XXXIII (Winter, 1956), p. 46.

<sup>&</sup>lt;sup>14</sup><u>Ibid</u>., p. 47.

<sup>15</sup>Ibid.

After a particular text has been chosen for study, the experimenter systematically deletes a random number of words. Taylor has found that deletion of every fifth word puts the blanks as far apart as they need be for most investigations of normal subjects. However, blanks may be farther apart if the experimenter feels it is necessary.

The mutilated texts are then reproduced with each missing word replaced by a standardized blank. Copies are distributed to every member of a test group. All subjects are asked to replace the deleted word. Each person is informed that all blanks are of uniform size and in no way indicate the length of the missing words.<sup>18</sup>

Every time a test-subject properly supplies the correct word, he scores one point. His "cloze" score for any text is the total number or percentage of deleted words he guesses correctly. For normal subjects, Taylor found nothing to be gained by accepting synonyms for the fill-ins. The problem of what an acceptable replacement for an intended word should be is obvious. He has found that only those words which exactly match the original should be counted.<sup>19</sup> However, in cases of dysphasia this might be dramatically impractical.

For the purposes of most studies it was found that the mutilation process need take no account of the kind or importance of specific words to delete. The results

```
18<u>Ibid.</u>, p. 416. <sup>19</sup><u>Ibid</u>.
```

based on the deletion of any word which gets systematically counted out appears to satisfy the needs for most investigations. In fact, Taylor feels this is a superior method. The limiting of missing words to nouns, verbs, adjectives, or prepositions is both unnecessary and awkward.<sup>20</sup> However, we might point out that here is an apparent potential for contextual study of ability of dysphasics to handle the various parts of speech.

Pilot studies also have shown that for most research the experimenter should have the text lengthy enough so as to permit a minimum of 50 word deletions. This is said to be necessary to represent a sensitive accounting of the difficulty of a passage. Enough words must be deleted to permit simple and more difficult ones to cancel each other in proportion to the frequency of their occurrence.<sup>21</sup>

The "cloze-procedure", as outlined above, has been, with slight modification, common to most previous applications of the concept. Only with some variation of these techniques could we foresee adaptation for formal testing of adult dysphasic patients. These variations are discussed below.

> <sup>20</sup><u>Ibid</u>., p. 420. <sup>21</sup><u>Ibid</u>., p. 419.

## HYPOTHESIS AND EXPERIMENTAL DESIGN

In our attempt to pioneer an application of the "cloze-procedure" to our subjects at hand with one of the many parameters of language behavior, we have randomly selected one specific area of possibility. Our question is: can we apply Taylor's concept to adults with dysphasia? More specifically; will "cloze" scores indicate existing knowledge of a given subject, and subsequently evaluate learning that has occurred after study?

<u>Re-statement of hypothesis</u>. Comparison of the results of pre-indoctrination and post-indoctrination scores obtained from a given text will indicate post scores are higher than pre scores; therefore, measuring the degree of learning that must have taken place.

Sample. "Cloze" tests were given to 15 adult dysphasic patients enrolled and/or discharged from the Crippled Children and Adults' Rehabilitation Center in Ft. Worth, Texas. It was determined beforehand that all test-subjects exhibited a basic language function. They ranged in age from 28 to 61 years with a median age of 45 years. Median time in school was 13 years, with 5 subjects having completed four years of college and one subject completing eight years of higher education. Three subjects did not go beyond the eighth grade.

Subjects were not screened for dysphasia subclassifications. Similarly, specific areas of disability such as anomia, agraphia, and alexia were not measured quantitatively. However, all patients did possess a basic language function as defined earlier.

<u>Test materials</u>. The stimulus material for this investigation consisted of five short texts on the same number of topics for each of the 15 test-subjects. (See appendix). Each of the five topics were graded at the fifth grade level of expectancy as to performance.<sup>22</sup>

All materials were typed in pica upper case letters on white  $8\frac{1}{2}$ " by ll" paper. No instructions were necessary as to procedure for taking the tests as all subjects had identical orientation sessions as to its mechanics by practicing other tests on different topics.

"Cloze" deletions were obtained by counting out five words and making a fifteen space blank for the sixth word. Five more words were typed, the sixth deleted, and so on, throughout all five passages. There was a total of 84 missing words for all passages combined.

There were four identical copies for each of the five texts. However, three were mutilated and one was not. The first copy was used for measuring each patient's pre-indoctrination aptitude. The second copy

22My Weekly Reader, January 18-22, 1960, pp. 1-4.

was used to measure post-indoctrination aptitude. The third copy was to measure the effect of a week's delay regarding recall of the subject matter. The unmutilated copy was used for indoctrination purposes.

<u>Procedure</u>. (Step One) On the first day of testing the text beginning -"On a stormy day..." (See appendix) was given individually to all test subjects in the same room by the same experimenter.

An alphabetical order of subject participation was used. That is, the person's last name falling nearest to the beginning of the alphabet was first.

The patient was issued one mutilated copy of the text. The only instruction given was as follows: "We will read the story together. I will stop when we come to a blank. Try to guess a word to fill each blank." As each attempt at "closure" was made, the investigator recorded the patient's response in the appropriate blank. A fifteen second time limit was imposed on the patient for each attempt at closure. If the subject failed to respond within this limit, the experimenter continued, leaving the space allocated for the missing word empty. Upon completion, the test was retained by the investigator and the subject was dismissed to the "waiting room".

This alphabetical order of individual participation was continued until all test subjects had completed their attempt at pre-indoctrination testing for this topic.

(Step Two) Immediately following the individual testing, the entire group of test subjects assembled and a 45 minute indoctrination period started: (Each indoctrination session for all five test days was identical in structure).

- A. The unmutilated texts were distributed.
- B. The experimenter read aloud the text.
- C. The patients read the passage in unison.
- D. Each member of the group was asked to paraphrase the contents of the article as briefly as possible.
- E. The group re-read the article in unison.
- F. The unmutilated forms were collected.

(Step Three) After the indoctrination session was concluded <u>Step One</u> was repeated. That is, each patient was tested individually using another mutilated copy of the text to measure post-indoctrination performance. The same order of subject participation was followed. The same time limits were imposed. Upon completion all test copies were retained by the experimenter and the subjects were dismissed for the day.

On the second day of testing the text beginning -"The quarter horse..." was used. The order of subject participation was reversed, so that, the person's last name falling nearest to the end of the alphabet was first. The remaining procedure was identical to that of the three steps described above for the first day of testing.

On the third day of testing the text beginning -"Food, as you know,..." was used. The person's last name which fell most nearly to the middle of the alphabet was first. Subsequent alphabetical order of participation was from this midpoint forward to the beginning of the alphabet and then from the end of the alphabet back to the midpoint. The remaining procedure was identical to the two preceding test days.

On the fourth test day the text beginning - "To a scientist..." was used. The order of subject participation and remaining procedure was identical to that used on the first day of testing.

On the fifth test day the text beginning - "Shrimp are found..." was the topic for the day. The order of subject participation and remaining procedure was the same as that followed on the second day of testing.

Obviously, immediate recall and learning were being measured through this "one day" pre and post indoctrination testing for each of the five texts.

How would this same group of dysphasic patients perform with delayed recall?

One week later Step One, discussed above, was re-initiated alone. That is, each day the same group took the same tests they had taken one week before. Steps Two and Three were omitted. At no time during this experiment were any of the test subjects allowed to take copies of the tests from the testing room.

<u>Scoring</u>. All testing sessions and scoring were supervised by the experimenter. Scoring was accomplished in the same manner as the previous application described above. One important modification was necessary, however. Due to the nature of the dysphasic's expressive inconsistencies, it was necessary to allow for appropriate synonym substitutions in lieu of the exact word replacement.

To determine if the synonym was an appropriate one, a panel of four speech therapists was gathered to agree as to each word choice when it differed from the word originally intended. Only those substitutions which were <u>unanimously</u> confirmed as appropriate were scored.

(Day)	First Day	Second Day	Third Day	Fourth Day	Fifth Day
Text	"On a stormy day"	"The quarter horse"	"Food, as you know"	"To a sci- entist"	"Shrimp are found"
(Sub- jects)	1 ) 2 3 4 5 6 7 8 9 10 11 12 13 14 15	15 14 13 12 11 10 9 8 7 6 5 4 3 2 1	8 7 6 5 4 3 2 1 15 14 13 12 11 10 9	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15	15 14 13 12 11 10 9 8 7 6 5 4 3 2 1

ORDER OF SUBJECT PARTICIPATION IN INDIVIDUAL TESTING

<sup>a</sup>The above numbers assume an alphabetical arrangement.

<sup>b</sup>For the second trial at closure each day, the order was the same. For the third trial, to measure delayed recall one week later, the subject participation and sequencing of the tests were identical.

#### RESULTS

Observation and generalization are extremely difficult when attempting to measure any parameter of linguistic ability in cases of dysphasic individuals. Additional influencing factors to this difficulty arise from the lack of selectivity of subjects and the "sheer numbers" of test subjects. The reader of this report must be aware that any discussion of results is limited by all these factors.

<u>Group success with "closure</u>." The data consisted of the responses of the 15 subjects to the 84 word deletions on five reading tasks. If all subjects had answered correctly all deletions there would have been a sum total of 1260 correct responses, or if each individual subject had answered correctly all deletions there would have been a sum total of 84 correct responses for each subject. All test scores combined yielded 435 correct responses or 34% of the deletions answered correctly for Trial One. Trial Two yielded 813 correct responses or 64% of the deletions answered correctly. Trial Three yielded 747 correct responses or 59% of the deletions answered correctly. In Table 2 are presented the mean number of correct responses made by the subjects on the three trials.

#### TABLE 2

MEAN NUMBER OF CORRECT RESPONSES ON THREE ATTEMPTS AT CLOSURE

	Trial l	Trial 2	Trial 3
Mean score (N=15)	29.1	57.5	49.8

A glance at the data indicating the differences between Trial One and Trial Two reveals a more than doubling of the individual scores, for this reason no test of significance was initiated. However, a mean gain was calculated and a standard deviation of the gains to indicate the variability of the scores was accomplished. Also, a standard error of the mean gain was computed to give an idea of the reliability. (See Table 3) The above scores should be taken only as tentative indications since the sample size was small and a normal curve had to be assumed without proof.

However, the significance of the results of Trial 3 was not as obvious. Therefore, a test of significance was accomplished.

In the selection of an appropriate statistical method the following points are of importance:

- 2. However, we were concerned with a statistical technique that is designed to predict the significance of the difference between two means obtained from the same test administered to the same group upon different occasions.
- 3. The study incorporated a sample size that was small and a normal curve had to be assumed without proof.

Specifically, our statistical approach needed to be adaptable to evaluate the effect of the intervening seven day delay upon the final scores; or to estimate the effect of no indoctrination or study interpolated between test and retest.

Garrett has outlined the procedure which meets our requirements most precisely. This method has been specifically designed for small groups which range in members from 12 to  $25.^{23}$  The method is basically a <u>t</u> test (the quantity <u>t</u> is the distance from the mean expressed in terms of standard error of the mean)<sup>24</sup> for small groups where the means are calculated.

Regarding the utilization of this technique, Garrett has offered this defense: "In problems like this,

<sup>24</sup><u>Ibid</u>. p. 191.

<sup>&</sup>lt;sup>23</sup>Henry E. Garrett, <u>Statistics in Psychology and</u> <u>Education</u> (New York: Longmans, Green and Company, 1947) pp. 210-211.

establishing and dealing with a mean gain or difference involves less calculation and is, also, to be preferred to the method of calculating standard errors (SE's) for each mean, an SE of the difference, and the correlation between the scores."<sup> $2_{0}$ </sup>

(1) Following this method, the mean difference between the subject's scores of Trial 2 and Trial 3 was established and a standard deviation or SD was calculated on these scores. This is given in the chart below.

## TABLE 3

MEAN DIFFERENCES, STANDARD DEVIATIONS, AND STANDARD ERROR REVEALED THROUGH THREE TRIALS AT CLOSURE

	<del>q</del>				Mean	Difference	SD	SE
Between	trial	1	\$	2		28.4	10	<b>2.</b> 56
Between	trial	2	38	3		7.73	6.40	1.65

(2) The next step, according to this technique, was to calculate the standard error or SE of the mean difference. The formula is:  $\frac{SD}{\sqrt{N}}$ . This gave a standard error of the mean difference which was 1.65.

(3) How significant was our mean difference of 7.73? This significance was expressed by the critical ratio or  $\underline{t}$  which became 7.73/1.65 = 4.68.

22

# 25<u>Ibid.</u>, p. 211.

(4) Each subject gave two scores, thus the degrees of freedom were (N-1) or 14. For 14 degrees of freedom using a table of <u>t</u> of 2.62, we will be exceeded in the positive direction in 1% of the trials. Since the critical ratio or <u>t</u> calculated above exceeds this number, the conclusion is that <u>the difference of 7.73 is highly significant beyond the .01 level of confidence.</u> Entries were taken from the .02 column of the <u>t</u> table for significant levels of .01 because the interest was in the probability of a positive difference as large as or larger than 7.73.<sup>26</sup>

## TABLE 4

## STANDARD ERROR AND t SCORE BETWEEN TRIALS TWO AND THREE

	SE	<u>t</u>	df (N-1)	Р
Between trials 2 and 3	1.65	4.68	14 <sup>a</sup>	01

awith 14 degrees of freedom a 2.62 is needed for a confidence level of .01 (one tail test). Therefore,  $\underline{t} = 4.68 = P$ \_\_\_\_01.

26 Ibid.

## CONCLUSIONS AND DISCUSSION

One's conclusions are always limited to the factors studied, and it would be impractical to draw generalizations which involve factors and conditions which have only been assumed. It would appear that the "cloze-procedure", within the conditions of this experiment, did significantly measure a degree of learning that must have taken place in the 15 cases studied. We might conclude, also, that this technique significantly predicted with sensitivity the effects of a one week delayed recall factor. However, the potential efficiency of this instrument is still undetermined; in that, this study did not incorporate a comparison with another standardized measure of learning. Undoubtedly, this is a possibility for future research.

Regardless of the results of a projected study of this nature, this investigator questions any implication that there is a measure of learning which is inherently superior, and that certain experimental conditions will permit this superiority to show itself, while other conditions might hide this superiority. However, we would like to suggest that this technique does have significant possibilities regarding the

measurement of learning with cases of adult dysphasia. Moreover, we would like to point out again that there is a significant advantage in Taylor's concept; in that, it seemingly taps simultaneously all of the complex determinants affecting the success of the communication act. Therefore, the multiplicity of applications this instrument seemingly possesses appears to be substantial.

"<u>Closure</u>" as a therapeutic technique. As reported above, success with "closure" seems to depend on most of the many factors which have significant influence on communication. This, therefore, is a "power" test of verbal performance. More appropriately, we might say that this is a <u>technique</u> which possesses a powerful <u>stimulation</u> potential in that it, also, stimulates simultaneously all of the complexities affecting word choice, both at the various levels of abstraction and through stretches in context.

The following application can be made: To take advantage of both visual and auditory stimulation in a "word finding" task, a topic of interest is chosen by the patient. A text of the topic is originated, and subsequently mutilated. The patient is asked to read aloud from the mutilated copy. The therapist reads in unison with the patient, stopping at each deleted word. The patient attempts to fill the blank. The patient has the advantage of being stimulated in a more normal manner by the sequential cueing of contextual material, has a topic of interest to him, and has both auditory and visual stimulation.

With some imagination various modifications of this application can be formulated. For example, the filling in of the deletions could become a writing exercise. By removing the visual stimulation of the printed text, a listening exercise is readily accomplished.

"<u>Closure</u>" as an evaluation instrument. Many of the areas of application utilized in therapeutics suggested above seemingly have almost equal potentials in the evaluation process. If the materials for measurement could be graded as to difficulty regarding levels of concreteness or abstraction, academic levels of achievement, and length; a rather comprehensive instrument for evaluation could be devised.

As we have reported earlier in the paper, the principle of "closure" has enjoyed success in indicating working intelligence, attention span, knowledge and learning of a subject, perception, language skill, listening comprehension, and reading comprehension with normal subjects. With the exception of measuring intelligence per se, we might hypothesize that "clozeprocedure" could become a useful tool in ascertaining

the other parameters of linguistic ability with dysphasic patients as listed.

Implications for future research. Obviously, any of the areas of study with normal persons discussed in this report could likewise be studied with dysphasic subjects. However, it appears to this writer that there are two problems reported in the literature with which this concept could be useful.

There are conflicting studies reporting the dysphasic's ability at "naming" tasks. Schuell suggests that small words are more difficult to elicit in contextual speech as they have a tendency to be overlooked and have slight auditory prominence. She has further implied that with increasing length of the stimulus unit more errors are found because of increased difficulty with reauditorization.<sup>27</sup>

With stimulation words presented in isolation, Siegel found a predominance of errors on long words.<sup>28</sup>

Another contradictory area has been reported by Siegel. Closely allied to the disagreement over the success or failure of the dysphasic to use "long" or

<sup>27</sup>Hildred Schuell, "Auditory Impairment in Aphasia: Significance and Retraining Techniques," Journal of Speech and Hearing Disorders, 18:1 (March, 1953), 14-21.

<sup>&</sup>lt;sup>28</sup>Gerald M. Siegel, "Dysphasic Speech Responses to Visual Word Stimuli," <u>Journal of Speech and Hearing</u> <u>Disorders</u>, 2:2 (June, 1959), pp. 152-160.

"short" words, is the dysphasic's ability to display equal performance with the various parts of speech in context.<sup>29</sup>

Regardless of these contrary reports, the "cloze-procedure" would appear to be an extremely useful instrument to facilitate further analyses of these problems.

In order to accomplish this task, the following basic plan is offered: (1) Originate a text of sufficient length so as to allow a gradual transition of the stimulus material from rather concrete to highly abstract. (2) Systematically delete every nth word as prescribed in previous application reported in this study. (3) However, repeat the deletion process often enough by rotating the order of word deletions so that eventually every word in the text has been withdrawn. For example, if every fifth word were chosen, the experiment would require five tests to complete the process. (4) A procedure for administering the test could be closely paralleled to the one we have utilized in this investigation.

29<sub>Ibid</sub>.

### SUMMARY

It was hypothesized that Wilson L. Taylor's "cloze-procedure" would measure learning exhibited through the performance of dysphasic adults.

The data consisted of the responses of 15 subjects to 84 word deletions on five reading tasks. Experimentation was initiated through three trials at "closure." Trial One was instigated to measure existing knowledge for the topics in the investigation. Trial Two was accomplished to test the ability of this technique to quantitatively measure a degree of learning that took place following indoctrination sessions. A third attempt was provided in order to ascertain the instrument's sensitivity to measure recall after the delay of a one week time factor.

An analysis of the results indicated a substantial increase in the raw scores and mean gain between Trials One and Two. However, a comparison between the scores of Trials Two and Three, was not as obvious; therefore, a statistical procedure was applied in the form of a one tail simple  $\underline{t}$  test. The results were highly significant beyond the .Ol level of confidence. It was concluded that "cloze-procedure" was successful in measuring learning within the conditions of this study, and that

the technique was sensitive to the degree of measuring delayed recall.

The multiplicity of applications this technique seemingly possesses was suggested. The concept of "closure" was discussed as a therapeutic aid, an evaluation instrument, and a potential resource for future research.

# APPENDIX A

# TABLE 5

## COMPARISON OF RAW SCORES OF 15 SUBJECTS IN THREE ATTEMPTS AT "CLOSURE" WITH NUMBER OF YEARS OF FORMAL EDUCATION

SUBJECTS	Years of Education	Trial l	Trial 2	Trial 3
1	8	13	34	28
2	20	61	74	63
3	8	40	58	37
4	12	18	60	44
5	12	22	56	52
6	16	28	66	62
7	8	16	39	21
8	12	19	45	45
9	16	25	59	63
10	16	31	71	58
11	12	18	48	38
12	12	26	60	56
13	12	69	77	77
14	16	24	55	50
15	16	27	61	53

## APPENDIX B

ON A STORMY DAY IN PHILADELPHIA, MORE THAN 200 YEARS AGO, LIGHTNING CRACKLED THROUGH THE SKY. A SMALL SILK KITE TOSSED ABOUT IN THE WIND. WITH THAT KITE AS "BAIT", A MAN WAS "FISHING" FOR LIGHTNING.

AS YOU KNOW, BEN FRANKLIN MADE A GOOD CATCH THAT DAY. ATTRACTED BY A WIRE ON THE KITE, LIGHTNING RAN DOWN THE WET KITE STRING TO A KEY. THEN FRANKLIN TRAPPED IT IN A JAR. HE PROVED THAT LIGHTNING WAS A FORM OF STATIC ELECTRICITY. THUS, ONE OF NATURE'S GREAT MYSTERIES WAS SOLVED. FOR CENTURIES, MAN HAD FEARED THE GREAT FLASHES OF LIGHT IN THE SKY. FRANKLIN USED HIS COMMON SENSE TO FIND OUT WHAT THESE FLASHES REALLY WERE.

THE QUARTER HORSE IS A HORSE THAT CAN RUN VERY FAST FOR A QUARTER MILE. BUT RACING IS NOT THE ONLY REASON FOR THE QUARTER HORSE'S POPULARITY.

THE QUARTER HORSE IS SEEN AT RODEOS AND HORSE SHOWS. IT IS THE FAVORITE MOUNT IN MANY RIDING CLUBS ALL OVER THE WEST. THE QUARTER HORSE HAS PROVED TO BE A GOOD "CUTTING" HORSE. CUTTING MEANS SEPARATING A CALF FROM THE HERD. THE CUTTING CONTEST IS BECOMING MORE AND MORE POPULAR. IN THE CUTTING CONTEST, A HIGHLY TRAINED HORSE, RIDDEN WITH A LOOSE REIN, CUTS A CALF AWAY AND HOLDS IT AWAY. HUMMINGBIRDS BURN FOOD SO RAPIDLY THAT THEY MUST EAT ALMOST CONSTANTLY. AS TWILIGHT COME, THE BIRDS STUFF THEMSELVES WITH INSECTS AND FLOWER NECTAR. THEN THEY SETTLE DOWN FOR THE NIGHT. THEIR BREATHING SLOWS AND THEIR BODY TEMPERATURE DROPS. BY THE MIDDLE OF THE NIGHT, THEY ARE HARDLY ABLE TO MOVE. THESE ARE SIGNS OF HIBERNATION. BY HIBERNATING FOR THE NIGHT, HUMMINGBIRDS ARE ABLE TO GO WITHOUT EATING FROM DUSK TO DAWN. TO A SCIENTIST, A CRANBERRY MAY NOT BE CALLED A CRANBERRY. IT IS A PLANT WITH SEVERAL SCIENTIFIC NAMES. A SCIENTIST DIVIDES PLANTS AND ANIMALS INTO DIFFERENT FAMILIES AND GIVES EACH MEMBER OF THE FAMILY A NAME.

A SCIENTIST KNOWS THAT THE CRANBERRY IS A COUSIN TO THE BLUEBERRY. BOTH OF THESE BELONG TO THE HEATH FAMILY. HEATH, OR HEATHER, IS A LOW SHRUB THAT GROWS ON THE MOORS (WASTELAND) IN GREAT BRITAIN.

REAL HEATHER IS NOT COMMON IN AMERICA. MANY PLANTS IN AMERICA, HOWEVER, BELONG TO THE HEATH FAMILY. THE BEAUTIFUL MOUNTAIN LAUREL AND TRAILING ARBUTUS ARE TWO MEMBERS OF THE HEATH FAMILY. SHRIMP ARE FOUND IN THE GULF OF MEXICO. MOST SHRIMP BOATS HAVE A CAPTAIN AND A CREW OF THREE. THE MEN MAY EAT AND SLEEP ON THE BOATS. SOME BOATS STAY OUT A WEEK OR LONGER. MOST OF THE SPACE IN THE BOATS IS KEPT FOR "SHRIMP PASSENGERS". THEY ARE CAUGHT IN A BIG NET. ON BOARD, THE SHRIMP ARE BEHEADED. ONLY THE TAILS ARE KEPT. THEY ARE PACKED IN ICE.

SHRIMP FISHING IS OFTEN DANGEROUS. STRANGE OCEAN CREATURES ARE SOMETIMES CAUGHT IN THE NETS. ONE CAPTAIN CAUGHT A SHARK LONGER THAN AN AUTO. THIS FISH HAS A LONG SNOUT WITH SAW-TOOTHED EDGES.

ON A S	FORMY DAY	PHILA	DELPHIA,
MORE THAN 200	ZEARS	LIGHTNI	NG CRACKLED
THROUGH THE SKI	Υ	SMALL SILK K	ITE TOSSED
ABOUT	THE WIND.	WITH THAT KIT	E
"BAIT", A MAN W	WAS "FISHING"		LIGHTNING.
AS YOU	KNOW, BEN	M	ADE A GOOD
CATCH THAT	•	ATTRACTED BY A	WIRE ON
	KITE, LIGH	TNING RAN DOWN	THE
KITE STRING TO	A KEY.	FRANKL	IN TRAPPED
IT IN A		PROVED THAT LIG	HTNING WAS
	FORM OF STATIC E	LECTRICITY. TH	US,
	OF NATURE'S GREA	I MYSTERIES WAS	
	FOR CENTURIES,	MAN HAD FEARED	
	GREAT FLASHES OF	LIGHT IN	
SKY. FRANKLIN	USED HIS COMMON		TO FIND

A QUARTER HORSE IS A THAT CAN
RUN VERY FAST A QUARTER MILE. BUT
RACING NOT THE ONLY REASON FOR
QUARTER HORSE'S POPULARITY.
THE QUARTER IS SEEN AT RODEOS AND
SHOWS. IT IS THE FAVORITE
IN MANY RIDING CLUBS ALL THE WEST.
THE QUARTER HORSE PROVED TO BE A GOOD
""HORSE. CUTTING MEANS SEPARATING A
FROM THE HERD. THE CUTTING
IS BECOMING MORE AND MORE IN THE
CUTTING CONTEST, A TRAINED HORSE, RIDDEN
WITH A REIN, CUTS A CALF AWAY

HOLDS IT AWAY.

## BIBLIOGRAPHY

## Books

Garrett, Henry E., <u>Statistics in Psychology and Education</u>. New York: Longmans, Green and Company, 1947.

Myklebust, Helmer R., <u>Handbook of Speech Pathology</u>. "Aphasia in Children: Language Development and Language Pathology, Edited by Lee E. Travis. New York: Appleton-Century-Crofts, Inc., 1957.

Articles and Periodicals

- Argill, M. H. "Modern Linguistics: Recovery from Aphasia," Journal of Speech and Hearing Disorders, 19 (December, 1954), 507.
- Schuell, Hildred. "Auditory Impairment in Aphasia: Significance and Retraining Techniques," Journal of Speech and Hearing Disorders, 18:1 (March, 1953), 14-21.
- Siegel, Gerald M. "Dysphasic Speech Responses to Visual Word Stimuli," Journal of Speech and Hearing Research, 2:2 (June, 1959), 152-160.
- Taylor, Wilson L. "Cloze-Procedure: A New Tool for Measuring Readability," <u>Journalism Quarterly</u>, XXX (Fall, 1953), 415-433.

"Recent Developments in the Use of 'Cloze-Procedure'," Journalism Quarterly, XXXIII (Winter, 1956), 46.

- "Cloze Readability Scores as Indices of Individual Differences in Comprehension and Aptitude," Journal of Applied Psychology, XLI (February, 1957) 19-26.
- Wood, Nancy E. "Helping the Aphasic Adult," Journal of Rehabilitation, 22:7, (January, 1956) p. 7.

Other Sources

- Taylor, Wilson L. "Application of 'Cloze' and Entropy Measures to the Study of Contextual Constraint in Samples of Continuous Prose," (Unpublished Ph.D. Dissertation, University of Illinois, 1954).
- . "Cloze Procedure," <u>Agrisearch</u>, 2:2 (February, 1956), 1-4.

"The Cloze Procedure: How it Predicts Comprehension and Intelligence of Military Personnel," <u>Technical</u> <u>Memorandum No. 13</u>, Urbana, Illinois: Division of Communications, University of Illinois, (1953).

My Weekly Reader, January 18-22, 1960, 1-4.

ROOM USE ONLY

JUN 25 1962 AR Pal
JUN 25 1962 52 Par
1111 12 1962
NOV S 1902
AUG
DEC-5-5-55 4
9961-2-9AV

•

NOR BOOK Binding General Book Binding Fort Worth, Texas

``

ŧ.

.

н С

