The entries in this annotated bibliography are divided into four major sections. The first includes some of the arguments pro and con on the cloze procedure and a few research studies along with some articles that give "how-to" recommendations. The second major section includes five subsections: the first deals with basic research concerning the validity and reliability of the cloze procedure as a device for assessing second language proficiency; the second includes a few references to cloze in the teaching of languages; the third is concerned with cloze as a measure of readability for ESL/EFL materials; the fourth treats cloze as an elicitation procedure for interlanguage data; and the fifth discusses oral/aural cloze research. The third major section includes references on cloze tests applied in languages other than English and also to social varieties of English. The last major section includes a subsection on general reviews of cloze research, followed by a subsection on reviews of cloze research applied to second language proficiency measurement. (Author/PMP)
RESEARCH WITH CLOZE PROCEDURE IN MEASURING THE
PROFICIENCY OF NON-NATIVE SPEAKERS OF ENGLISH:
AN ANNOTATED BIBLIOGRAPHY

John E. Oiler, Jr.
University of New Mexico

CAL*ERIC/CLL Series on Languages and Linguistics
Number 13

ERIC Clearinghouse on Languages and Linguistics
Center for Applied Linguistics
1611 North Kent Street
Arlington, Virginia 22209

April 1975
I. Pros and Cons on the Basic Procedure--Research and Debate

II. Research with Cloze Procedure--Especially with Non-Native Speakers of English
   A. As a Testing Device
   B. As a Teaching Procedure
   C. As a Measure of Readability
   D. As an Elicitation Device
   E. Aural/Oral Cloze Test Research

III. Research with Languages Other Than English or Social Varieties of English

IV. Reviews
   A. General
   B. Language Testing and Cloze Procedure
Havelock Ellis once said that every artist writes his own autobiography. Perhaps he would draw the line at bibliography. However, even a bibliography reflects the author's bias in its selectivity. If it is annotated, as this one is, the bias is apt to be even more apparent. Since the flurry of research with the cloze procedure which began in the late 1960s and which has continued even to the time of this writing in 1975, the cloze procedure has become a classroom notion, if not a household word. Many teachers of language, including English as a second or foreign language, and bilingual educators have begun to talk about and experiment with the cloze technique. It is probably true, nonetheless, that the technique has been vastly underexploited to date. For this reason, it is hoped that the research cited here will provide guidelines for further research and experimentation, as well as applications.

Typically, the procedure has been used for measuring the proficiency of students in a classroom situation, or in an institutional testing context; it has sometimes been used for the investigation of certain learner or material characteristics. With respect to learner characteristics, the technique is eminently applicable to the study of so-called errors or interlanguage phenomena, yet only a handful of such studies can be cited in this bibliography. The cloze procedure provides a ready basis for focusing on particular sorts of grammatical rules and characterizing either learner grammars or the effects of teaching strategies. Concerning material characteristics, the cloze procedure has often been used to deal with questions about readability, or the difficulty of aurally presented material.

There are many ways of constructing and administering cloze tests, and there are methods of combining the cloze procedure with other testing or data elicitation procedures, e.g., dictation (see the references to S. Johansson below). However, most studies have utilized the standard procedure of deleting every nth word (usually every 5th, 6th, or 7th) from a passage of written prose. The alternative of deleting selected parts of speech, or "content" words as opposed to "function" words, etc., has also been used,
but less frequently. The possible combinations of the two procedures and the various possibilities within each are nearly endless. Nevertheless, all cloze studies have one feature in common: they require a subject or group of subjects to replace missing verbal material in some context.

The bibliography which follows is divided into four major sections. The first includes some of the arguments pro and con and a few research studies along with some articles that give "how-to" recommendations concerning cloze procedure. The heavy debate is represented principally in the articles by Rand and myself in that section.

The second major section includes five subsections: the first deals with basic research concerning the validity and reliability of cloze procedure as a device for assessing second language proficiency; the second includes a few references to cloze in the teaching of languages; the third is concerned with cloze as a measure of readability for ESL/EFL materials; the fourth treats cloze as an elicitation procedure for interlanguage data; and the fifth discusses oral/aural cloze research.

The third major section includes references on cloze tests applied in languages other than English and also to social varieties of English.

The last major section includes a subsection on general reviews of cloze research, followed by a subsection on reviews of cloze research applied to second language proficiency measurement.

I. Pros and Cons on the Basic Procedure—Research and Debate


Recommended a blank-in-sentence procedure for teaching and testing ESL/EFL. Includes a recommendation concerning cloze procedure.


A test which enables users to equate different groups in terms of their respective reading abilities. Utilizes a mathematical formula for determining readability coefficients of reading selections.

Bowen, J. Donald. TESOL research for the classroom. TESOL Quarterly, 1972, 6, 551-62.

Reviews a number of research projects conducted at the University of California at Los Angeles. Includes a number of references to cloze test research.

Clark recommends that the cloze procedure be validated against more direct measures.


Proposes the possibility that cloze procedure might be sensitive to register differences.


Criticizes standardized tests of language proficiency including CILT, ALGU, and TOEFL. Suggests a "contextualized" task-based test which the authors believe will be superior to cloze procedure and other techniques.


Criticizes the contextually acceptable method of scoring--it is not well defined, according to the authors. The cloze test employed contained 38 blanks and was correlated with the various part scores on the Michigan Test of English Language Proficiency, Forms F and H. The study used 105 subjects, but occasionally reports results with groups as small as 8. No significant differences were found between exact-word and acceptable-word scoring methods. (The fact that the results here contradict Stubbs and Tucker, 1974, suggests a replication.)


Reviews the literature on language testing in relationship to findings in linguistics and psycholinguistics--particularly in first and second language acquisition studies. Argues for a more comprehensive theory of linguistic competence in order to explain language test data. Insists on tests that are related to the context of communication in ways that are as natural as possible. Mentions findings concerning the cloze procedure and other integrative tests.


Discusses fill-in-blank procedures and multiple-choice versions of language tests.
Ingram, Elizabeth. The psycholinguistic basis. In Spolsky, B. (Ed.), Current Trends in Language Testing, in preparation. Presents a number of definitions of the term psycholinguistics and relates them to practical interpretations of language testing, especially cloze procedure. Will be of interest primarily as a thorough review of the theoretical literature. Highlights some research findings, but makes no attempt at completeness.


Films were interrupted and 92 subjects were tested on their capacity to supply missing information. The technique seems applicable to the measurement of differences in subject ability in comprehending films. (Possible applications to second language situations where audio-visual techniques are used are obvious.)


Posits a mathematical relationship between the estimated redundancies of passages of English prose. The estimates are based on independent samples of the same population of subjects. Claims that there is a linear relationship between the logarithms of estimated redundancies of cloze passages tested against independent samples of children at corresponding grade levels. Concludes that "the present experiments suggest that meaningful estimates of information rate, behaviorally derived by cloze test procedures may prove as feasible—and potentially as productive—as have estimated sample variances of obtained test scores" (p. 38).

———. Development of a standardized reading test designed to discriminate effectively at the adolescent level. Mimeograph, University of Saskatchewan, Saskatoon, Canada, 1973.

Reports the development of cloze-type tests suitable for children with reading ages ranging up to ten-and-a-half years. Eminently applicable to second language testing problems.

Moss, Kenyon. Cloze testing in reading comprehension. TESL Reporter, 1972, 6, 3.

Reports research on the use of the cloze procedure to evaluate reading comprehension of 7th and 8th graders. Includes a brief bibliography on the cloze procedure and second language testing.


Cites the high degree of variance overlap between cloze tests, essay writing, oral interview, various forms of paraphrase recognition, and completion tasks as evidence in favor of an underlying conceptual dependency network of some sort which might be
characterized as a grammar of expectancy. Argues that the investigation of language tests, like cloze procedure, might provide much useful information as an elicitation device for the investigation of developing grammars at various stages.


Provides a brief historical overview of the development of pragmatics and develops the notion of a pragmatic expectancy grammar in terms of the interpretation of an exemplary sentence. A primitive example of a fragment of an expectancy grammar provides a model in terms of which to explain performance on cloze tests and other integrative language tests.

-----. Expectancy for successive elements: Key ingredient to language use. Language Annals, 1974, 7, 443-452.

Discusses the notion of an expectancy grammar as a basis for explaining the variance which is common to many integrative tests, including cloze procedure.


A critique of the article by Rand entitled "Integrative and Discrete-Point Tests at UCLA." In particular this paper reinterprets Rand's factor analysis of cloze procedure and a number of other tests in the Connolly study.


Argues that integrative tests, like cloze procedure, dictation, translation, and certain types of multiple-choice items, such as sentence paraphrase or selecting the central idea of a paragraph, are superior to discrete-point tests which focus on phonemic contrasts, particular morphological endings, and the like.


Outlines techniques for constructing, administering, and evaluating cloze scores for ESL purposes. Cites correspondences between frustrational, instructional, and independent reading levels.


Presents an argument against blank filling (a special variety of cloze procedure) and in favor of translation.


Argues that earlier research using cloze procedure and dictation failed to establish the superiority of integrative tests.
such as cloze and dictation, over discrete-point tests (actually, multiple-choice tests of reading, grammar, and vocabulary). Rand reports a factor analysis from the Connolly, Masters thesis, which used a four-scale aural interview, in addition to four cloze tests, four dictation passages, and two composition items (a sentence completion task and a paragraph completion task). Rand found that one factor accounted for 70% of the variance and received loadings from all 21 of the variables included.

----- Some evidence for the predictive validity of the ESLPE.

Reports on a preliminary validation of the ESLPE Winter 1974 Version (W74). The test includes the following subsections:

- a 15-point multiple-choice grammar section focusing on typical foreign student difficulties;
- a 15-point grammar section concentrating on "non-standard dialect problems";
- a 25-blank cloze test (items "ranged from five to eight words apart" and were selected by the test constructor);
- a 20-point reading test consisting of multiple-choice items on three passages;
- a 25-point listening test involving multiple-choice items based on a short lecture and a ten-line dialogue; and two dictations worth 15 points each, consisting of about 125 words each.

Rand concludes that the test is a good placement device correlating highly with teacher ratings of students tested.


Proposes cloze procedure along with the noise test and a number of other integrative tests as measures of "overall language proficiency." This paper is one of the early arguments in favor of what is currently being called "global language proficiency testing."


The first published statement concerning the "cloze procedure" developed by Wilson L. Taylor. Originally the procedure was conceived as a readability measure for passages of prose. In this study, evidence and arguments are presented for the superiority of a measure like cloze procedure over traditional readability formulas such as the Flesch or Dale-Chall formulas.


Reports on research employing cloze procedures for listening comprehension tests and also suggests the possibility of using the
cloze procedure for testing foreign language proficiency. An excellent seminal article.


II. Research with Cloze Procedure—Especially with Non-Native Speakers of English

A. As a Testing Device


Fifty-item tests achieved reliabilities of .95 and better. Concludes that cloze technique is a more sensitive measuring instrument than the judgments of reading experts. There were also several independent checks of validity which indicated cloze scores are valid measures of reading comprehension ability for non-native speakers of English. This is a thorough, milestone research project in the evaluation of cloze procedure and its applications in the teaching of English as a foreign language.


Another report on blank length. Also investigates the feasibility of having the students write all answers in a column on an answer sheet in the margin. Revealed that both versions of cloze procedure correlated well with the Watts Sentence Test. Actually the test using exact-length blanks correlated slightly better with the Watts Sentence Test than did the test using standard-length blanks. Subjects were non-native speakers of English in the fourth, fifth, and sixth grades in Papua-New Guinea schools.
A report of research on comprehension in reading. Paper presented at the Third International Reading Association World Congress on Reading, Sydney, Australia, August, 1970. ERIC ED 015 325.

Reported on the use of blanks of the word deleted. Cloze tests constructed by this procedure worked as well as cloze tests constructed by the standard procedure of using a blank of uniform length throughout. A practical implication is that teachers will be able to construct cloze tests by using photocopying facilities. Used non-native speakers of English.


A descriptive statement of the procedure and a very brief attempt to describe a model. The model harks back to Taylor (1953), who claimed the technique was a good device for testing the goodness of fit between a speaker or writer's system and a reader or listener's system. Reports that blanks of the same length as the deleted word work as well as blanks of a standard length. Claims there is no difference between ranking of passages or subjects for the two blank procedures.

Application of a cloze procedure to English learned as a foreign language in the Territory of Papua and New Guinea. English Language Teaching, 1972, 27, 67-72.

Explores several different scoring procedures: exact word, words that make sense in the context, and words that are of the same grammatical class as the deleted word. The result is that the relative ranks of the students and tests remain about the same, regardless of the procedure used. This replicates findings with native speakers. See also Oller, 1972, on the same topic.


A pioneering study in the application of cloze procedure to the measurement of English as a second language and also bilingual dominance. Although this study did not incorporate validity checks in terms of additional tests, comparisons between native and non-native performance on the English cloze tests revealed expected differences, and scores in English were generally correlated with course grades (rank orders were checked). Scores in Amharic seemed less systematic (perhaps due to different scoring criteria).


Correlates the UCLA ESLPE (which includes a section of cloze) with judgments of nativeness.

Briefly mentions why the cloze techniques were retained in Cohen's follow-up research and further development of the UCLA ESLPE. Discusses the rationale behind inclusion of various parts on the UCLA, dating from the fall of 1972.


Describes a variety of reading-item types, including cloze procedure, for the measurement of proficiency in a second or foreign language. Provides a fairly thorough review of the literature.


Cross-validates an oral interview procedure against cloze tests and dictation.


Surveys literature on the use of cloze procedure, especially in the teaching or testing of English as a foreign language. Reports data from a correlation of an experimental cloze test with the UCLA English as a Second Language Placement Examination, Form 2C (summarized and expanded statistically by Oller and Conrad, 1971). The performance of the cloze procedure as a discriminator between different ability levels was also checked against the performance of students at various levels in the English as a Second Language Program at UCLA and at a junior college in the Los Angeles area.


Cloze procedure differentiated among Mexican-Americans-dominant in English, dominant in Spanish, and Anglo groups. Crawford compared the cloze tests results against other tests including the Stanford Achievement Test and Comprehensive Tests of Basic Skills Level 2, Form Q. Some subjects at both third- and sixth-grade levels also took the Gilmore Oral Reading Test, Form C. The cloze tests proved to be valid and reliable with both groups of subjects.

Darnell, Donald K. The development of an English language proficiency test of foreign students using the clozentropy procedure. Speech Monographs, 1970, 37, 36-46.

Compares performance on a cloze test against the Test of English as a Foreign Language for 48 foreign students at the University of Colorado. The cloze test was also administered to
200 native speakers of English at the same University. The cloze test was about as reliable as the TOEFL (.86). The correlation between the two tests (total scores) was .83. An analysis of variance revealed that the content and difficulty of test material, major subject area of examinees, and level and major of the native comparison group had significant influences on the index of English proficiency. Scoring was based on the degree of conformity of non-native responses to the norms established on the basis of the 200 native speakers. This meant it had to be done by computer, which makes the technique somewhat cumbersome for classroom purposes. The highest correlation between the cloze test and any part of the TOEFL was with listening comprehension.

Davies, A. Two tests of speeded reading. Paper presented at the Language Testing Symposium, Georgetown University, Washington, D.C., March 1:74. Failed to replicate certain earlier findings with cloze procedure. However, Davies used a test construction method that relied on experimenter judgments rather than standard every nth word deletion.

Estrada, F.X. The effect of increasing syntactic complexity on reading comprehension. Unpublished M.A. thesis, University of Southern California, 1969. Investigated the performance of 75 fourth-grade Navajos and native speakers of English on cloze tests of differing levels of difficulty. Results indicated that cloze test difficulty is significantly related to the complexity of sentence structures included. The native English speakers out-performed the Navajos on all levels.

Evans, Gordon, and Haastrup, Kirsten. Experiments with cloze procedure. Department of English, Danmarks Laererrojskole, Endrupborg, 1973. Submitted for publication. Discusses the development of proficiency tests in English by a Nordic test-development group set up in 1971. The article "deals with two of the subtests, both of which use a modified cloze procedure" (p. 3). The test finally included was the result of pre-testing in which items with low discrimination indexes were eliminated. The final draft of the cloze portion included 17 items with an average item discrimination of .53. The cloze test portion requires 9 out of the 51 minutes devoted to the written section of the test. Also, a partial dictation of the sort used by Johansson is included in the listening comprehension section.

Goldmann, Tamar T. English as second language testing and the prediction of academic success. Unpublished M.A. thesis, University of California, Los Angeles, 1972. Workpapers in TESL: UCLA, 1972, 6, 97 (Abstract). Compared correlations between part scores and the total score of the English as a Second Language Placement Examination at UCLA against grade point average. The correlations were between .08 and .23--too slight to be of much predictive value.

Reports the development of cloze tests for elementary school children in grades 2-7 in African primary schools. The tests were too difficult for the second-grade level, but gave satisfactory results at higher levels. Item-to-total correlations ranged between .85 and .95. Individual item reliabilities ranged from .60 to .91. Hofman concludes that the procedure is generally applicable at the elementary grade levels, but the particular tests which he devised worked best with students in the fourth through seventh grades.


Endeavored to establish the factor validity of use of cloze tests with students for whom English is a foreign language. Two hundred and eighty-four elementary school pupils of standards four, five, and six completed cloze tests with every eighth word deleted. Multiple-choice comprehension questions were made over the same material to test vocabulary, knowledge of stated facts, grasp of organization, perceiving causality, identifying the main theme, and making inferences. One factor accounted for 63 percent of the variance in the matrix, a factor called 'reading comprehension.' This study supported those of Bormuth (1966) and Anderson (1969). The cloze procedure measures skills closely related to or identical to those measured by a multiple choice comprehension question test (Riley, 1973, p. 21).


Four questions were asked: (1) Is a cloze test of prepositions feasible, reliable, and valid? (2) Are response distributions for native and non-native speakers similar or different? (3) Will contrastive analysis enable the researcher to predict responses from a group of students from the same native language background? (4) Will the test provide useful diagnostic information concerning trouble spots for non-native speakers in general? All questions were answered in the affirmative except for (3). Only a small percentage of errors were predictable on the basis of contrastive analysis.

Irvine, Patricia; Atai, Parvin; and Oller, J.W., Jr. Cloze, dictation, and the Test of English as a Foreign Language. Language Learning, 1974, 24, 245-52.

The Test of English as a Foreign Language was administered to 159 non-native speakers of English in Iran. They also took a cloze test and two dictations. Part scores and total scores were intercorrelated between the TOEFL and the two experimental measures.
Results confirm earlier findings. The cloze and dictation correlated better with the listening comprehension part of TOEFL than any other subsection of that examination. The authors conclude that the data provide strong support for integrative testing procedures.


The experimenter administered the Certificate of Primary Education Examination in mathematics, English, and general knowledge plus cloze tests in English. Subjects were students in Kenyan schools who did not speak English as a native language. Correlations between the cloze tests and the CPE English examination revealed "concurrent validity."


155 foreign students in ESL classes at UCLA in the fall quarter of 1971 and 44 ESL students in university extension served as subjects for this study. Eighty-five of the students in the regular ESL classes practiced taking either a cloze test or dictations throughout the quarter. Pre-test, post-test results indicated that there was no difference in the improvement rate of the experimental groups as compared to that of the control groups. A follow-up experiment with students at the lower level in extension classes indicated that there may be some improvement when cloze passages and dictations are focused on particular "phonetic and grammatical points."


Compares the performance of Japanese students at the secondary level who have experienced English as a foreign language in the elementary school against those who have not. Three cloze tests were administered. A response frequency analysis for each item on all three tests is included along with other statistics.


Advances the hypothesis that the high correlations among tests of listening comprehension, speaking, reading and writing are an indication of test validity. Discusses correlations between cloze tests and dictation.

-----. Scoring methods and difficulty levels for cloze tests of proficiency in English as a second language. Modern Language Journal, 1972, 56, 151-158.

Investigates five different scoring methods and concludes a slight superiority of an acceptable-word scoring method over the
exact-word scoring method. The acceptable-word method counts as correct any response which does not violate any of the constraints of the passage. Item analyses, intercorrelation matrices, as well as first, second, and third order partial correlations are provided for cloze tests compared with the UCLA English as a Second Language Placement Examination, Form 2A Revised. Reliabilities for the three cloze tests included in the study ranged from .93 to .99 on the Kuder-Richardson formula 20 for subjects, and from .80 to .95 on the same formula for items. The study utilized 398 subjects. There was a tendency for cloze tests to correlate better with the sections of the UCLA ESLPE which were more integrative in nature.


Compares correlations between dictation tests and cloze tests and the UCLA English as a Second Language Placement Examination, Form 2A Revised, and dictation correlated with UCLA ESLPE, Form 1 only.


A cloze test scored by the exact-word procedure is correlated with the UCLA ESLPE, Form 2C. The multiple correlation between all sections of the ESLPE and the cloze test was .88. Surprisingly, the cloze test correlated best with the dictation subsection on the UCLA ESLPE. This replicated a similar finding of Darnell, 1968, with the TOEFL.


Every other preposition was deleted from a passage of prose in English. The test was taken by 19 native speakers of English, 53 native speakers of Turkish, and 110 foreign students entering UCLA from a wide variety of language backgrounds. Item analyses and response frequency data are reported. Certain intralingual confusions were exhibited, regardless of language background.


Reports results from an experiment using cloze procedure to measure the proficiency of Japanese secondary school students, some of whom had had training in English as a foreign language in the elementary grades, and some of whom had not. Non-FLES students apparently are able to overtake FLES students by the eleventh grade.


This 114-page report constitutes the most thorough validity
study on the Test of English as a Foreign Language ever undertaken. It investigated correlations between sections and sub-sections of the TOEFL, as well as intercorrelations with various experimental tests, including a paragraph rewriting task developed by Kellog Hunt, a standard cloze test (open-ended), several multiple-choice sub-tests (including a multiple-choice version of a cloze test), an essay, and an oral interview. The results, on the whole, provide additional support for integrative over discrete-point tests, and replicate a number of findings using the cloze procedure.


Urban and rural Thai children, grades two to five, 151 boys and 142 girls, were the subjects in this study. Raven's Progressive Matrices, class grades in reading, and average academic grades were used as assessment measures. Passages from texts in use were selected from grade levels one to six and mutilated by an every-fifth-word deletion. Subjects completed cloze tests at their own grade level and the next level above and below. Exact-word scoring was used. Results showed that cloze passages were ranked consistently among subjects, indicating that cloze assesses readability. Cloze scores increased with grade level and ability, indicating that cloze is sensitive to differences in reading ability. Cloze scores correlated positively with school achievement and scores on Raven's Progressive Matrices, indicating the validity of cloze tests for assessing Thai elementary school children (Riley, 1973, p. 25).


Develops an argument for global testing procedures. Mentions the cloze technique as modified in the Spolsky et al. noise test.


Investigated oral cloze procedure and part scores on TOEFL. Finds no evidence of unique variance that can be associated with the Listening Comprehension sub-test which is not also associated with Reading Comprehension and other sub-tests.


Correlates the American University at Beirut English Entrance Examination with a cloze test scored by exact-word and acceptable-word methods. Concludes that the cloze test in general is a

B. As a Teaching Procedure


Discusses some of the psychological processes underlying reading skill. Recommends the use of cloze procedure as a testing device to measure readability and comprehension and also as a teaching procedure which will hopefully force "the hesitant student to think, to attack his reading problems actively by drawing on his several kinds of knowledge about English" (p. 182).


Claims validity for cloze both as a measure of comprehension and as a teaching device--though cloze was not more effective than other techniques--and compares cloze with Metropolitan Achievement Test Reading Section and with multiple-choice tests of reading comprehension.


Discusses the reading process and argues why the cloze procedure should be a suitable device for enriching the learning experience of students in ESL/EFL classes.


Used cloze procedure in helping to teach German function words. The German words were inserted in a passage of English prose at appropriate points. After the training program, students were able to translate correctly 60 percent of the German words.

C. As a Measure of Readability


Describes the use of cloze procedure "as a measure of reading difficulty with second language learners" and "how teachers may construct their own cloze tests, administer and score these, and interpret the cloze scores for grading reading materials" (p. 36). A very readable and helpful "how-to" article.

Discusses the development of a reading laboratory modeled after the Scientific Research Associates reading laboratories. Used the cloze procedure to grade 210 selections in terms of reading difficulty.


Investigated the correspondence of cloze test scores with "traditional levels of reading--independent level, instructional level, and frustration level." Reveals that 53, 44 to 53, and less than 44 percent correspond respectively to the three preceding levels. This was determined with cloze tests scored in a standard way--counting only exact-word responses as correct.


Reported satisfactory results of grading materials with cloze procedure. Concludes that the exact-word scoring method is preferable (for reasons of economy of time and effort) for purposes of estimating readability.


Cloze tests were administered to two classes of 35 pupils each at the elementary and junior high levels in Papua-New Guinea. Results indicated sensitivity to such subtle variables as style of writing, degree of conformity to children's beliefs, and interest value of the materials.

D. As an Elicitation Device


Used a cloze-like procedure for the elicitation of data from 9- to 11-year-old Anglo children and from native Spanish-speaking adults. The study concentrates on questions about the usefulness of the cloze procedure as an elicitation device and also presents a thorough analysis of the rating of responses according to their
acceptability by a group of adult native speakers of English. Provides insight into the bridge between interlanguage analysis and proficiency testing.

See also the Johansson papers under section II.E.; Oller, Bowen, Dien, and Mason under section III; and Oller and Inal under section II.A.

E. Aural/Oral Cloze Test Research (with native and non-native speakers)


Discusses "theoretical and practical considerations in the development and use of aural proficiency tests" (p. 1). Reviews research with "direct" and "indirect" techniques of language proficiency measurement. Clark considers the cloze procedure an indirect measure.


Revealed that an aural cloze procedure used with children was capable of discriminating among sociolinguistic backgrounds for Anglo, Spanish, Black, and Navajo subjects.


Used two aural cloze tests where messages recorded on tape deleted a word every five seconds by means of random masking noise. The cloze tests were reliable at .70 and .80 and intercorrelated at .73. Attitude toward the message (one "persuasive," the other "expository") proved to be insignificantly related to either cloze tests or to the multiple-choice tests of comprehension used.


Replicates earlier findings with a noise test reported by Spolsky and challenged by Johansson. Also replicates findings of Oller and Johansson with straight dictation and partial dictation (that is, dictation combined with cloze procedure). Recommends the noise test, oral cloze tests, and straight dictation as measures of second language proficiency.

Compares the performance of an aural cloze procedure using every-fifth-word deletions from a tape-recorded message (white noise was substituted for blanks) against the Brown-Carlsen Comprehension Test, Form BM, and California Reading Test, Advanced, Form Y. Concludes that cloze procedure is a valid and reliable measure of listening comprehension. The aural cloze reliability was .92 and .93 for subjects at USC and Cambridge Adult School, respectively.


A slightly expanded and revised version of the paper presented in Copenhagen, 1972 (see below).

-----. Controlled distortion as a language testing tool. Paper presented at the Third International Congress of Applied Linguistics, Copenhagen, Denmark, August 1972.

Criticizes the Spolsky noise test. Claims that partial dictation is superior to the noise technique.


Utilizes a testing procedure that combines dictation with cloze procedure. The student reads a passage with blanks while he hears the complete passage dictated. The task is to fill in the blanks. This procedure was recommended by Brooks (1965), although Johansson developed it independently on the basis of a study of dictation at UCLA and some of the research on cloze procedure listed in this bibliography. Johansson provides a useful and interesting analysis of errors and their types.

Kaplan, Robert B., and Jones, R.A. A cloze procedure test of listening for university level students of ESL. Mimeographed University of Southern California, 1970.

Explores the possibility of using an oral version of cloze procedure at the university level. Conclusions were not very encouraging.


Among the presentation methods were techniques using a preview and repetition. 311 students were tested at a junior high in Tucson, Arizona. Passage type, presentation, method, and deletion all proved to be significant factors by an analysis of variance.
III. Research with Languages Other Than English or Social Varieties of English


Five passages in Malay were ranked by reading specialists and teachers. Every fifth word was deleted, with five versions over each passage so that each word in every passage was deleted in at least one of the versions. Pupils at grades four, five, and six were tested. The tests had high re-test reliabilities and discriminated well among subjects.


Samples of contextual test forms. Reports on the application of a modified cloze procedure to tests in English, French, Spanish, Chinese, Russian, and Portuguese. The research was sponsored by the Department of Defense and focused on the testing of aptitude, achievement in language courses, and overall language proficiency.

See also Bowen, 1969, on Amharic.


The study was designed to investigate the feasibility of using cloze-type items for College Board foreign language tests. Reviews the literature and discusses the development of cloze tests in English, French, and German. Test characteristics, including reliabilities and validity measures, are reported. An experiment on the possibility of aural cloze is also included. It is noteworthy that the test preparation did not follow standard cloze procedure of deleting every nth word from a passage of prose. Rather, certain items were selected on the basis of pre-testing. Concludes that the cloze technique measures language skill quite indirectly and that there is very likely a specific skill in taking cloze tests which is not correlated very highly with language proficiency. The study has two major limitations: first, it
utilized a very small sample for many of its generalizations; and second, instead of using a proficiency measure as a validating basis, the study used the MLAT. The exact-word scoring method was used.


Used a cloze test in English and in Hebrew which deleted every sixth word from passages in the respective languages. The intent of the authors was to use language proficiency as a criterion against which various attitude measures could be compared. No results are reported in this paper, however.


The study concludes that cloze tests in French as a second language have the same applicability as in English.


Reports good success with cloze procedure in Vietnamese. Cloze scores are apparently sensitive to the quality of translations.


This study parallels the Anderson and Rahman study discussed above. Five passages were selected, and five different versions of cloze tests were constructed over each passage by deleting every fifth word until every word was systematically deleted in each passage. Cloze rankings agreed well with rankings by experienced teachers. The tests discriminated between teachers and correlated with language examinations. Counting acceptable-word responses and exact-word responses in scoring yielded an intercorrelation of .98. It was concluded that cloze is as "suitable a measure of reading comprehension in Chinese as in European languages" (Riley, 1973, p. 25).

Compares performance of children in five languages (Czech, Polish, French, Spanish, and English) and concludes that the cloze procedure is in fact "translatable." Comparable results can be achieved across languages. See also Oiler, Bowen, Dien, and Mason, 1972.


Reports detailed analysis of cloze tests in English and Thai. Compares cloze test scores with grades from students' standardized final English exams, and with scores of a group of 35 American high school seniors who took the English versions of the tests. This study was coordinated with the thesis by Dien completed in the same year. Reports detailed response frequency analysis for items on the English cloze tests.


Summary of Mason's M.A. thesis, listed above. This report also includes full copies of the tests in English and Thai.


Among other things this study investigated types of errors made by native and non-native speakers of English. Mean scores for native and non-native speakers on translated passages indicated rough comparability between the original passage and its translation. Native speakers made fewer responses which did not conform to at least some of the contextual constraints in the passages.

Potter, T.C. Reading Comprehension among Minority Groups: Child Generated Instructional Materials. 1968, ERIC ED 031 540.

Stories generated by third-grade children who spoke Black English were more readable, as judged by the cloze measure, than the Standard English stories commonly used in the classroom.


Mentions cloze procedure along with several other testing techniques which might be useful in fulfilling the goal stated in the title.

Schaier, Barbara; Cooper, Robert L.; and Fishman, Joshua A. Language, technology, and persuasion: A Middle Eastern example. Research Committee on Sociolinguistics, Language in Science

Reports findings from the Cooper and Fishman article listed above. Does not mention any results from cloze procedure. The major result was that Hebrew turned out to be more effective for arguing in scientific terms, while Arabic was more effective for persuasion in "traditional messages."


Stendahl, Christina. The relative proficiency in their native language and in English shown by Swedish students of English at university level. English Institute, Gothenburg University, 1972. Used the cloze procedure to explore the relationship between native and foreign language proficiency in Swedish and English respectively. Found a surprisingly high correlation between students' proficiency in the two languages. That is, students who tended to score well on tests in Swedish also tended to do well on the tests in English, and vice-versa.

Wijnstra, Johan M., and Van Wageningen, Nico. The cloze procedure as a measure of first and second language proficiency. Submitted for publication (Dutch Journal of Psychology), 1973. Used cloze procedure tests in Dutch and Spanish. A multiple-choice version of one of the tests in Dutch was constructed. The result was that the multiple-choice form correlated at .85 with the open-ended standard cloze procedure scored by the acceptable-word method as recommended by Ollier (1972). Replicates a number of findings from previous studies.

IV. Reviews

A. General


Boyce, Max W. A comprehensive bibliography of the cloze procedure. State College of Victoria at Toorak, Victoria, Australia, 1974.


Suggest that cloze procedure might be used to measure second language proficiency.


Provides an excellent review of research relating cloze scores to information theory and derived estimates of redundancy. Defines an uncertainty reduction index in terms of cloze test scores which, among other things, can be used for estimating an individual's reading ability. Suggests the possibility of applying uncertainty reduction index estimates in the investigation of the translatability of cloze tests across languages. Implies that the same estimates of uncertainty reduction could be applied to the English as a foreign language context.


B. Language Testing and Cloze Procedure
See especially the annotated bibliography by Riley, below.


This is probably the most comprehensive bibliography on the topic in existence.


Reviews results with the Spolsky et al. test since 1969. Concludes that more research is needed, but the method remains substantially useful. Gradman suggests its use as an elicitation device for the investigation of error types common to "a particular level of language proficiency" (p. 47).


This report mentions several studies in Europe that are using some version of cloze procedure. Several of the reports mentioned are not published.


This volume includes papers presented at the Language Testing Symposium held in Washington, D.C., March 1974, at Georgetown
University. Items of interest are referred to separately in the present bibliography.

Kennedy, Graeme D. The testing of comprehension. In Spolsky, B. (Ed.), Current Trends in Language Testing, in preparation. This 127-page manuscript is a monumental contribution to the understanding of listening comprehension and reading comprehension. An excellent and very thoroughgoing review of the research is included. Naturally, the subject is broad and the review cannot include all references, but it is remarkably complete. Kennedy also takes account of research with cloze procedures and dictation along with a great many other testing procedures which he capably describes.

Kiare, George R. Assessing readability. Reading Research Quarterly, 1974-75, 10, 62-102. Excellent review of the research on cloze as applied to readability. Includes recent work with non-native speakers of English and with languages other than English.


----- Language testing. In H. Douglas Brown and Ronald Wardhaugh (Eds.), Survey of Applied Linguistics. Ann Arbor: University of Michigan, in press. An introductory survey of language testing research and findings. Traces the roots of discrete-point testing to contrastive analysis theory and structural linguistics. Develops a definition for the notion "efficiency of the learner's internalized expectancy grammar." Cites research with cloze procedure and other integrative techniques as well as error analysis studies which support the proposed definition.

----- Pragmatic language testing. Language Sciences, 1973, 28, 7-12. Introduces some notions from a pragmatic theory of language and extrapolates their relevance to practices in language testing. Argues that cloze procedure is very closely related to other integrative language tests, such as oral interview, dictation, and reading tests of various sorts.

A non-technical introduction to a number of testing procedures, including discrete-point tests on grammatical structure, sentence interpretation, word order, vocabulary, subject-verb agreement, structural parallelism, case of pronouns, comparison of adjectives, formation of adverbs, and formation of irregular verbs. Also discusses integrative tests of reading skills, including vocabulary in context, sentence paraphrasing, identifying the central idea of paragraph, cloze procedure dictation, and composition.


Surveys briefly the history of the development of the Test of English as a Foreign Language and research relating to it. Correlations between the TOEFL and cloze tests are reported.


Includes sections on review articles, general methodology and rationale, readability, teaching English as a second language, languages other than English, and other miscellaneous studies. Riley's bibliography is considerably broader in scope and less focused than Robinson's bibliography (cited in section IV.A. above).


Discusses the history of language testing and argues in favor of integrative tests like cloze procedure. Stresses the need for tests that are sensitive to sociolinguistic variables.
OTHER TITLES IN THIS SERIES

1. Preparing and Using Self-Teaching Units for Foreign Languages,

2. A Selected Bibliography on Bilingual/Bicultural Education,

3. A Selected Bibliography on Language Teaching and Learning,
   by Sophia A. Behrens and Kathleen McLane, 1975.
   Available from EDRS: ED 100 189.

4. A Guide to Organizing Short-Term Study Abroad Programs,

5. Working Papers in Linguistics,
   by Tim Shopen, 1975. Available soon from EDRS.

6. A Selected Bibliography on Mexican American and Native American Bilingual Education in the Southwest,
   Available soon from EDRS.

7. Using Community Resources in Foreign Language Teaching,

8. A Selected Bibliography of Films and Videotapes on Foreign Language Teacher Training,
   by Peter A. Eddy, 1975. Available soon from EDRS.

9. ERIC Documents on Foreign Language Teaching and Linguistics:
   List Number 13,
   by Peter A. Eddy, 1975. Available soon from EDRS.

10. Effects of Social Situation on Language Use: Theory and Application,
    by William Check, Theodore B. Kalivoda, and Genelle Morain, 1975. Available soon from EDRS.

11. Radio in Foreign Language Education,

12. Error Analysis in the Classroom,
    by Patricia B. Powell, 1975. Available soon from EDRS.
ERIC (Educational Resources Information Center) is a nationwide network of information centers, each responsible for a given educational level or field of study. ERIC is supported by the National Institute of Education of the U.S. Department of Health, Education, and Welfare. The basic objective of ERIC is to make current developments in educational research, instruction, and personnel preparation more readily accessible to educators and members of related professions.

The ERIC Clearinghouse on Languages and Linguistics (ERIC/CLL), one of the specialized clearinghouses in the ERIC system, is operated by the Center for Applied Linguistics. ERIC/CLL is specifically responsible for the collection and dissemination of information in the general area of research and application in languages, linguistics, and language teaching and learning.

In addition to processing information, ERIC/CLL is also involved in information synthesis and analysis. The Clearinghouse commissions recognized authorities in languages and linguistics to write analyses of the current issues in their areas of specialty. The resultant documents, intended for use by educators and researchers, are published under the title CAL-ERIC/CLL Series on Languages and Linguistics. The series includes practical guides for classroom teachers, extensive state-of-the-art papers, and selected bibliographies.

The material in this publication was prepared pursuant to a contract with the National Institute of Education, U.S. Department of Health, Education, and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their judgment in professional and technical matters. Prior to publication, the manuscript was submitted to the Teachers of English to Speakers of Other Languages for critical review and determination of professional competence. This publication has met such standards. Points of view or opinions, however, do not necessarily represent the official view or opinions of either the Teachers of English to Speakers of Other Languages or the National Institute of Education.

This publication will be announced in the ERIC monthly abstract journal Resources in Education (RIE) and will be available from the ERIC Document Reproduction Service, Computer Microfilm International Corporation, P.O. Box 190, Arlington, Virginia 22210. See RIE for ordering information and ED number.

For further information on the ERIC system, ERIC/CLL, and the CAL-ERIC/CLL information series, write to ERIC Clearinghouse on Languages and Linguistics, Center for Applied Linguistics, 1611 North Kent Street, Arlington, Virginia 22209.