A Reexamination of Cloze Procedure in the Teaching and Testing of English as a Foreign Language

William D. Patterson

The basic premise of Cloze Procedure is that a native English reader of a cloze passage should be able to fill in (usually single) deleted lexical items created by a consistent deletion process. The filling of the deleted items, furthermore, depends on how well the reader can supply universal or global understanding among the concepts in a given string of discourse. The higher the incidents of exact replacements, the “better” the student understands the passage. The purpose of this research is to examine the recent history of the Cloze Procedure in teaching and testing, mainly among non-native speakers of English, and to assess the efficacy of Cloze Procedure within the pedagogy of English as a Foreign Language.

Cloze Procedure has had a substantial history in its occurrences in the literature of teaching of English as a Second/Foreign Language as well as in the fields of psychology, business and native language education. Its earliest modern recorded use was as a completion test designed by Ebbinghaus (1897) to measure intelligence. Though the term that Ebbinghaus used was not ‘Cloze’, it was his contention that what would later be called Gestalt ‘Closure’ would be a situation in which:

“pupils get prose text with gaps (e.g., missing syllables, missing parts of syllables, and missing words) and are asked to fill in gaps in a meaningful way.” The number of syllables and parts of syllables that need to be filled in is indicated for each gap (Wilhelm, n.d.).

Later (and independently of Ebbinghaus) Taylor developed completion tests as educational tools in which gap-filling exercises were used to predict the readability of prose passages. Taylor reported that the results seemed to better the measurement of difficulty levels of text among native English speakers (Taylor 1953). Incidentally, Taylor is often incorrectly credited with devising the Cloze Procedure when in fact the term Cloze derives from the Gestalt ‘closure’ (http://en.wikipedia.org/wiki/Cloze_test). In a review of the literature of Cloze Procedure, Walter describes it as:
[a] method of interrupting a message from a “transmitter” (Writer or speaker), mutilating its language by deleting parts, and so administering it to “receivers” (Readers or listeners) so that their attempt to make the patterns whole again potentially yield a considerable number of cloze units. (Walter, 1974).

Taylor’s research dealt with native speakers of English and his method was not primarily used as an EFL teaching or testing tool. Taylor was more concerned with using the Cloze Procedure for establishing readability factors. As was said above, Taylor’s conclusion was that Cloze was a better predictor of reading difficulty than other empirical measurements. (ibid.) Chatel expands the use of close from monolingual assessment to foreign language teaching but agrees with Taylor as to the interpretation of the results of cloze when she states, “The Cloze Procedure is a multifunctional strategy, which can be used for reading diagnosis and reading comprehension.” Some of her research dealt with non-native English speakers.

She expands upon the purpose of Taylor’s use of Cloze Procedure and posits the use of Cloze as an instructional medium by adding, “Diagnostically, it is an effective way to determine how a reader is utilizing the context of a sentence or paragraph to get the meaning of the text. In addition, once the diagnostic function of a cloze test is done, the same text can be used instructionally making [the Cloze Procedure] a truly multipurpose strategy.

As a further caution, Chatel warns against the use of Cloze as a diagnostic tool until “students are familiar with the procedure”. Otherwise, test takers tend to focus on the blank where the word occurred and try to fill in a suitable answer based on immediate context. Correct use of Cloze requires that students pay attention to the whole passage and text, not just to the words surrounding a blank. Therefore, if Cloze Procedure is to be used effectively, students must be taught how to work with a cloze passage.

This procedural outline is three-fold: first students should read the entire passage; second, be instructed to fill in every blank, and last, reread the passage one last time to make any minor readjustments to the filled in gaps.

As early as 1978 Radice stated the efficacy of cloze testing in the field of ESL (italicized instructions mine):

1. Ease of Preparation: Cloze tests are easy to prepare, with the direction that a coherent passage be chosen. Common materials have been textbooks, essays and magazine articles. Preserve the first and last sentences of a passage and delete every (preferably) 7th word, replacing it with a line 12 spaces long, numbering the blanks for a total of 50 items.
2. Ease of Administration: Instruct students to read the entire passage one time; go back and fill in one word per blank, and then read the passage one last time. Allow at least 10 minutes. No dictionaries should be allowed.

3. Immediate knowledge of results: For strict correction, match the students’ responses with the original deleted word. Accept only exact replacements.

4. Feedback to the teacher: The instructor can have immediate feedback as to the suitability of the material for the group, i.e., low percentage of correct replacements vs. high preventative. The lower the percentage of correct replacements the more difficult the material is for the students. This is practical for determining if a group of students could comprehend a text-book, news magazine, or book of essays.

5. Suitability for group corrections: Depending on how many students in a group incorrectly answer collocations; parts of speech or idioms, etc., the teacher can address the problem to the entire class.

6. Flexibility: The cloze procedure can be used in a variety of situations. It can also be constructed with lexical deletions, multiple-choice alternatives, near replacements, words in the same grammatical category, etc.

7. Separation of grammatical difficulties: As in the lexical deletion mode, the test maker can delete singular/plural nouns, present, past tenses, modals, prepositions, articles and so on as demanded by the training or teaching goals to be accomplished.

**CLOZE TESTS THEN AND NOW**

Research suggests that a cloze passage be approximately 250-500 (depending on the deletion frequency to be used) words in length with the first and last sentences of the passage left untouched (no deletions should occur in those sentences). Commencing with the second sentence of the paragraph, every nth word is removed and replaced with a standard sized, numbered blank line. Cloze Procedure purists such as Oller and Carrol insist that a word interval of every 7th word seems to be the optimum deletion interval. Furthermore, if a 7th deletion falls on a proper name that could not be guessed even from context, the deletion is mandatory.  

Oller and Carrol also contend that a deletion rate of every seventh word calls up more of a global understanding of the passage. A more frequent deletion rate (every fifth word) has less dependence on global discourse and more on the collocation of nearby elements. In other words, deleted items can be more easily reconstructed from immediate clues. Nor does an interval of every 10th word seem justified. It is the maintenance of the deletion of every 7th word
that seems to bring about just the right level of global reasoning and restoration of deleted items on the part of the test taker and thereby give the best results for “understanding” the connections within a Cloze passage.

Another rule in cloze test construction is that it is not permissible to shift a deletion to a position before or after where it would normally occur, thus making a string of, say, six, a Cloze deletion, and then a string of eight words of text. This tends to violate a global constraint on the written discourse.

Prior to Taylor’s use of the Cloze Procedure lexical deletion was the prevailing method of creating cloze tests. Lexical deletion involves the removal of specific items at a non-determined interval. Taylor (1957) maintained that a consistent (every n\textsuperscript{th}) deletion was “superior to lexical deletion in efficiency of construction as well as equal or superior in its correlation to ability measures.” Taylor also found the Cloze Procedure a more reliable indicator of difficulty of written materials than standard formulas such as the Flesch Reading Ease or the Fog Index. (1953)

It seems obvious that the deletion of every n\textsuperscript{th} word is mechanically and administratively simple in terms of test construction. For example, if a test maker desires to test the occurrence of prepositions in a verb/particle construction, all the test writer has to do is delete the prepositions within a passage. Similarly, to test article usage the test writer inserts blanks where the articles occur (or to make the task more demanding, the test writer could insert blanks before all nouns in the passage to ascertain if students can determine where no article is used). However, once the prepositions and particles are removed the global dynamics of the discourse are lost and the exercise becomes a mere reconstruction of phrasal prepositions.

This (fixed interval deletion) procedure has been shown by Taylor (1957) to be superior to lexical deletion (in which words are deleted based on their relationship to the text) in efficiency of construction as well as equal or superior in its correlation to readability measures. Cloze tests in Taylor’s study included a mean of 41 blanks. Participants were instructed to read through the passage and fill in the blanks with the words that best fit into the sentence. If participants did not know the exact words that fit in the blank, they were instructed to guess. Participants were given 10 min to complete each cloze test.

1) I have found that in my own administrations of cloze tests, that if one of the deletions should fall on a name or number (such as a date, score, age, etc.) it is permissible to start the initial deletion at one or two words prior or subsequent to the first deletion instance and then maintain the 7th deletion criteria from that point, thus avoiding an “impossible to supply from context” answer.
In a more recent study (Fanguy 2004), the Cloze procedure was administered, not to students of English as a Foreign Language, but to customers of large American chain stores to discover “as an alternative, [if] the Cloze test provides a means to measure how “average” consumers actually understand key sections of written material.” (Fanguy 2004) In fact, Fanguy et al., have what is to date the only web based Cloze Procedure this writer has found. Most of the classic cloze tests are print based.

To judge the readability and understandability of corporate privacy statements, a web-based (computer generated) cloze passage was shown to test takers. The software automatically disregarded the first and last sentences of the passage and then deleted every 5th word. An example of an actual test before the test taker hit the ‘submit answers’ button is shown in Figure 1.

The tests that Fanguy were generated for four US companies in order to determine how much of their privacy statements were understood by customers. Surprisingly, only one of four
tests displayed a comprehension rate of more than 60%. As much of the research heretofore published sets a level of around 60% correct answers in a close is the minimum requirement for unassisted reading, Fanguy’s research shows that at least three out of the four passages used were not easily understood by native speakers. Figure 2 shows the results of the study.

From Fangway (2001)

**GRADING AND SCORING OF CLOZE TESTS**

Since the beginning of the use of the cloze test as a measurement of language proficiency, discussion has ranged between two opposites, viz., exact replacement and varying degrees of acceptable answers. Those who espouse the strict replacement method usually cite ease of test construction, administration and grading. Carroll (1959) suggested that for a non-native teacher who is less than perfectly bilingual, construction and scoring is more or less mechanical and provides a relatively easy way to measure the level of a student’s competence in a foreign language.

Whichever of the two methods is used, i.e., exact replacements of deleted words or replacements that are considered acceptable depending on the context of the passage and the grammatical constraints of the utterance, it would seem that for the sake of simplicity, exact replacement is the most consistent and that accepting substitutes can be subjective and unreflective of linguistic competence. However, Carroll (1959) gives four rules for scoring non-exact matches: “(1) the word substituted did not change the meaning of the text, (2) the word had to be consistent with facts, (3) the word had to be the same part of speech or a valid substitute, and (4) grammar was considered only for articles”. It is obvious that the judgments call for a high degree of language sophistication and are beyond the grasp of most
FL instructors.

Misspellings may be allowed in the strict replacement model if the word is recognizable, but this is outside the present study.

In another study, Folse (2006) tested 152 students who were in the USA in preparation for college courses. In his research with Cloze Procedure, Folse employed modified cloze passages of short sentences for the purpose of helping the students learn and retain lexical items. It is fair to say that Folse was concerned more with vocabulary retention than with the global learning that Cloze Procedure engenders. However, this shows the flexibility and utility of gap-filling exercises.

Moving into the latter half of the 20th Century and into the field of English as a Foreign/Second Language, there are those who advocate its use and those who remain unconvinced of its usefulness. Alderson (1979), for example, claimed that testing readability via cloze testing is valid with native speakers but has limited effectiveness among non-native speakers. Fotas (1991) claimed cloze procedure as a viable testing instrument.

**CONCLUSION**

Even though many would believe that the jury is still out on the efficacy of Cloze Procedure as a testing and teaching tool in the EFL classroom, there is ample evidence that Cloze does, in fact measure global language comprehension skills, much the same as those skills were measured in native speakers of English who underwent Cloze Procedure testing.

In spite of the fact that the cloze procedure has prompted much discussion and criticism, there is ample evidence that it does measure reading comprehension and higher order processing. There are to date viable studies that support the Cloze Procedure (Jonz 1990, Farhady and Keramati 1996). Their notion of text based design dealing with NPs promises to enhance the validity of cloze as more than a testing device in which the focus is on collocation or immediate contextual clues. Incidentally, Farhady and Keramati used both types of scoring with the NP replacement and the result was a correlation of 0.899.

**APPENDIX**

A commonly used test is offered by the University of Michigan and included herein show what a modern test is. This is a commercial endeavor and should be taken as such. The first criticism of the test is the frequency of deletions ranging between 10 and 20. This is a lexical deletion test and may well be designed to encourage students to buy it online or to be used for a beginning to intermediate level student.

The University of Michigan test, surprisingly, is not a strict replacement test but a multiple
choice test “…designed to give you a taste of the cloze section in any Michigan Proficiency test. If you can get eight out of ten correct, (or better) you’re in pretty good shape for the exam. However, if you get less than eight correct, I would recommend more practice.”

Cloze Test Practice

Vaccination is a term coined by Edward Jenner for the __1__ of administering live, albeit weakened, microbes to patients, with the intent of conferring immunity against a targeted form of a related __2__ agent. Vaccination (Latin: vacca—cow) is so named because the first vaccine was derived from a virus affecting cows: the cowpox virus, a relatively benign virus that, in its __3__ form, provides a degree of immunity to smallpox, a contagious and __4__ disease. In common speech, ‘vaccination’ and ‘immunization’ generally have the same colloquial meaning.

Vaccination efforts were initially met with some resistance before early success brought widespread acceptance and mass vaccination campaigns were undertaken. The eradication of smallpox is considered the most spectacular success of vaccination. The last natural case of smallpox was discovered on October 26, 1977 in Somalia. This date is considered the __5__ of smallpox.

Modern-day critics of vaccination are concerned that __6__ vaccination plays a role in autoimmune disease and autism, though large-scale scientific studies have failed to find a link.

In the generic sense, the process of triggering immune response, in an effort to protect against infectious disease, works by ‘priming’ the immune system with an ‘immunogen’. Stimulating immune response, via use of an infectious agent, is known as immunization. Vaccinations __7__ the administration of one or more immunogens, in the form of live, but weakened (attenuated) infectious agents, which __8__ are either weaker, but closely-related species (as with smallpox and cowpox), or strains weakened by some process. In such cases, an immunogen is called a vaccine.

Some modern vaccines are administered after the patient already has contracted a disease, as in the cases of experimental AIDS, cancer and Alzheimer’s disease vaccines. Vaccinia given after exposure to smallpox, within the first four days, is reported to __9__ the disease considerably, as vaccination within the first week is thought to be beneficial to a degree. The essential theory behind such immunizations is that the vaccine triggers __10__ immune response than the natural infection itself.

(Choices for cloze test)

1. process transmission procedure growing
2. sick unhealthy disease sickness
3. healthy demi sick weakened
4. deathable dead deadly deathly
5. eradication existence killing terminate
6. childhood adulthood childish children
7. produce involve need acquaint
8. virtually normally usually always
9. attenuate diminish eradicate disappear
10. super huge enormous superior

(http://www.michigan-proficiency-exams.com/cloze-test.html)

REFERENCES


A Reexamination of Cloze Procedure in the Teaching and Testing of English as a Foreign Language


