

MULTIPLE-CHOICE AND CLOZE TESTS: WHAT DO THEY REALLY MEASURE?

Miriam Sester Retorta

Centro Federal de Educação Tecnológica do Paraná

The objective of this article is to show the results of a MA research carried out from July 1995 to August 1996. The aim was to find out what mental process(es) a testee goes through, while taking a multiple-choice and cloze test, in order to arrive at a correct answer. Cohen (1984:70-1) states that "there is a small but growing literature on how students go through the process of taking language tests." Very little has been done in the field of language testing and even less has been done in second-language testing. The conclusions reached in this research is just a tiny light at the end of a long tunnel which can bring insights for further research in this area.

Although we have been in a communicative era of language teaching/learning, the multiple-choice and cloze tests, which have been considered instruments of measurement of the psychometric-structuralist and psycholinguistic-sociolinguistic era, are still used in Brazil to measure reading comprehension. However, since little is known in terms of a psycholinguistic perspective these tests are often questioned if they are efficient instruments to measure reading comprehension.

Henning (1987) wisely claims that any kind of measurement is bound to be misused. One might have problems if there is a defect in the scale, or if the product being weighed might be altered due to other variables. Language testing, which is also an instrument of measurement, is not different. The problem of test measurement error is particularly relevant in relation to the reading skill, because there is no obvious measurable product. The reading skill differs from the writing and speaking skills in that we cannot perceive a reader's performance whereas the performance of writers and speakers is fully observable. This makes measurement of the reading skill a difficult issue.

In order to seriously investigate any test-method it is necessary to understand in depth the construct involved the skill which is being tested, which in our case is reading comprehension. Therefore, before actually exposing the research, a brief history of the different views of reading process must be taken into consideration.

The reading process

Before the 60's and 70's, reading was known as being a bottom-up process. The reader was a passive decodifier of graphemes. His task was to build up the meaning, which was crystalized in the text, putting together letter by letter, word by word, sentence by sentence. Therefore, a reading problem was necessarily a decodification problem. A good reader was the one who dominated the structure of the language which was of central importance.

In the mid-60's and early 70's, the concept a top-down process emerged. Reading was

seen as a "guessing game"^[1]: a reader, who was seen as an active agent in the reading process, would make predictions, create hypothesis or expectations about the text based on prior knowledge and check the text for confirmation or refutation of these expectations. Language knowledge was of secondary importance.

Therefore the reading concept switched from a precise process, which involved exact, detailed, sequential perception and identification of letters, words, spelling patterns and large language units to a selective process, which involved partial use of available minimal language cues selected from perceptual input on the basis of the reader's knowledge of the subject/world. As information was processed, tentative decisions were made to confirm, refute or refine the hypothesis.

From the 80's to the current moment, reading is seen to be an interactive process where both bottom-up and top-down processes are of central importance. Eskey (1989:94), defines this process as follows:

the interactive model does not presuppose the primacy of top-down processing skills - the gradual replacing of painful word-by-word decoding with educated guessing based on minimal visual cues - but rather posits a constant interaction between bottom-up and top-down processing in reading, each source of information contributing to a comprehensive reconstruction of the meaning of a text.

In this model the usage of knowledge of the language is as important as the usage of knowledge of the subject/world to create hypothesis and expectations about the text. The reader continues being seen as an active agent in the reading process and the pure decodification of the bottom-up process gives room for the automaticity theory (see Eskey 1989:??). The concept of the interactive process has also been changing and currently theorists like _____ prefer to include to the bottom-up, top-down interaction other elements such as text-type knowledge.

TESTING ISSUES

The concept of testing has also changed throughout the years. In an early stage, during World War II testing was, according to Weir (1990:02), in a psychometric-structuralist era. Baker (1989:35) observes that "from the structuralist descriptions of language theorists took the hierarchical analysis of language for the purpose of teaching and testing. From psychometric testing theory it borrowed the discrete-point 'objective' test formats and the statistical apparatus used for test development." Since discrete point tests were considered inadequate as indicators of language proficiency, another kind of test gained support: the integrative test (Weir, 1990:03). Testing then entered a new phase: the psycholinguistic-sociolinguistic era. This approach viewed integrative testing as an instrument which measured the ability to unite disparate language skills in ways which more closely approximated the actual process of language use. Oller (1986: 37), who is a faithful defender of integrative tests, states that:

The concept of an integrative test was born in contrast to the definition of a discrete point test. If discrete items take language skill apart, integrative tests put it back together. Whereas discrete items attempt to test knowledge of language one bit at a time, integrative tests attempt to assess a learner's capacity to use many bits all at the same time, and possibly while exercising several presumed components of a grammatical system, and perhaps more than one of the traditionally recognized skills or aspects of skills.

However, there have been many theorists who are critical of integrative tests. Weir (1990:5-6) supports the idea that the cloze test and dictation do not elicit spontaneous production from the testee, and the language norms used are not the testee's, but the examiner's. He also claims that both types of tests measure knowledge of the language system and not the ability to operate it in a real-life situation. In other words, these tests tell us about a candidate's linguistic competence and not about his performance ability.

Alderson (1980:59) questions the reliability of cloze tests. Based on the results of his research he showed that testees' scores on a cloze test are affected by altering the point where the deletions are started from, or by using a different n/th rate deletion.

Davies (1981:182), also points out the deficiencies of integrative tests:

Although Oller has claimed that his integrative tests represent total language proficiency better than any other single test or combination of tests, this is not in itself an argument in favor of the unitary competence hypothesis, as measures such as cloze and dictation are so integrative that they contain most or all language abilities anyway. High correlations between cloze and other measures may only reflect that they are measuring different skills which are highly correlated among individuals; however, this does not mean that there will be no individuals whose performances in the various skills differ considerably.

The controversies which arose from discrete-point and integrative tests gave way to a new approach to language testing: communicative testing and therefore a new era emerged - the communicative era.

According to Bachman (1991), the 1980's and 90's could be characterized as decades of 'communicative testing'. There is a shift in emphasis from the linguistic to the communicative dimension. The emphasis is no longer on linguistic accuracy, but on the ability to function effectively through language in particular contexts or situations. Weir (1990:9) demonstrates where communicative testing stands by stating:

Communicative testing, as well as being concerned with what the learner knows about the form of the language and about how to use it appropriately in contexts of use (competence), must also deal with the extent to which the learner is actually able to demonstrate this knowledge in a meaningful communicative situation (performance), i.e. what he can do with the language - his ability to communicate with ease and effect in specified sociolinguistic settings.

However, this new approach is not immune from measurement defects. A major problem the test writer faces when constructing and applying a communicative test is lack of reliability. Although this kind of test has high validity, it lacks reliability. Weir (1990:15) observes:

Other than serious marker reliability problems, associated with the assessment of performance, the major issue affecting an adoption of a 'communicative' approach to language testing is the generalisability of the results produced by a test.

Since communicative testing presents reliability problems, some test designers/ teachers prefer to continue using the multiple-choice and cloze test, which, a priori, are reliable

instruments to measure reading comprehension. However, although they might be extremely reliable, they have low construct validity. This problem was detected in a research carried out from november 1995 to august 1996 at the Universidade Federal do Paraná in order to the obtain a master's degree.

RESEARCH METHODOLOGY

The subjects

Twenty-three Brazilian informants who have the Cambridge Proficiency Certificate were asked to take three multiple-choice tests and one cloze test, and to answer two questionnaires - one for each type of test. The informants did not have any previous knowledge about the purpose of the study.

It was decided that only people who had the Cambridge Proficiency Certificate would be subjects of the study because we wanted to avoid a problem that might arise if this variable were not controlled. The inventory of strategies used in multiple-choice and cloze tests would be useless if one could argue that the strategies used by the testees were due to poor reading skills. Only proficient English speakers, readers, writers and listeners pass this examination and therefore, only efficient readers took the tests.

The methodological approach which was used in this dissertation was basically similar to that used by Cohen (1984). His data was collected through a verbal self-report, whereas in this research the data was collected through written self-report.

The tests

The multiple-choice tests

Three multiple-choice tests were chosen from 'the University of Cambridge International Examinations'- Certificate of Proficiency in English - June 16, 1993 (see appendix 2 p.101). The first text (see appendix 2 p.102) - passage one (406 words) - was about the relationship between an old man and his child; the second text (see appendix 2 p.104) - passage two (333 words) - was about the writing process; the third text (see appendix 2 p.106) - passage three (351 words) - was about TV advertisements. Each text was followed by 5 multiple-choice questions, each question containing 4 alternatives of which only one was the correct answer. By the end of the three multiple-choice tests each informant had therefore answered a total of 15 questions.

Some undesirable variables were controlled as follows:

- 1) We made sure that the topics of the tests were familiar to all the informants, i.e. the content was part of the testees' knowledge of the world. This avoided a situation where testees made mistakes not because they did not know how to read, but because they did not know enough about the subject.
- 2) The tests were taken from the Cambridge Proficiency Examinations because they are valid and reliable tests. Any multiple-choice test, before officially being used as a test, needs to go through a validation process which consists of pre-testing (the test maker prepares a test and

has it piloted); item analysis, (the test maker analyses each alternative correcting any faults found). The Cambridge Proficiency Examination goes through a thorough validation process before being used throughout the world, which is the reason why these tests were chosen.

The cloze test

For the cloze test we used a text selected from the Cambridge Proficiency Examinations. However, the deletion procedure was altered: the Cambridge Proficiency test followed a rational deletion criterion also called 'variable-ratio method' (Oller 1986:345). According to Weir (1990:48) this criterion is used when items are selected for deletion based on what is known about language, about text difficulty and about the way language works in a particular text. Since the intention of this research was to deal with the fixed-ratio method (Oller - 1986:345), i.e. when the deletion rate is mechanically set, usually between every fifth and eleventh word, we decided to keep the text and change the rational criterion to fixed-ratio deletion, deleting every 6th word. Twenty words were deleted. The first sentence as well as the last two were left intact so that the testee could have a better idea of what the text was about. The topic was about 'Illiteracy' (219 words), which can be assumed to be part of the testees' knowledge of the world (see appendix 03 p.108). The informants were instructed to fill in the blanks with only one word and to read the whole text before starting to answer it. Since none of the texts were on topics unfamiliar to the testees we assume that they were not extremely difficult, such that the reader could not understand the message.

The questionnaire

Together with the tests a separate handout (see appendix 1 p.92) was designed with the following information:

a) Instructions about what was expected from the informants, i.e. they would have to read the texts, answer the questions in the multiple-choice tests and fill in the blanks in the cloze. As soon as they finished each test, they had to go to the handout and write about how they arrived at each answer, i.e. the mental process they went through in order to answer the questions or fill in the blanks.

b) Directly after the instructions, there was a guideline containing questions which showed the subjects some possible procedures someone may adopt when reading, and some other procedures a test-wise candidate may adopt when taking a test. They were asked to use the guideline just for reference: they were not supposed to answer the questions. Both the instructions and guidelines came before the space provided for the testees to write their comments on each item. There was one set of instructions and guidelines for the multiple-choice tests and another set for the cloze test. This part was read out loud together with the informants so they could ask any questions about the task they would have to perform. They were also instructed not to stop in the middle of a test - they had to take one test and soon afterwards write about each question. Breaks between tests were allowed.

Criteria adopted for scoring the tests

Criterion for the multiple-choice tests

Since there was only one correct answer for each question, the criterion was to consider correct the answers which were officially chosen by the Cambridge Proficiency Examination board.

Criterion adopted for the cloze test

The informants' performance in the cloze test was measured not by the exact word method but by the acceptable word method, i.e. any word which was grammatically, semantically and stylistically suitable was considered correct.[\[2\]](#)

Data analysis and identification of strategies

The answers from the open questionnaires from both tests were analysed as follows:

1. The answers were classified according to whether they were correct or incorrect.
2. Frequency distributions of correct and incorrect answers for the two test methods were established.
3. Procedures used by subjects were identified and listed.
4. The frequency of occurrence of each procedure was measured.
5. These procedures were then classified into two main categories: reading micro-skills and test-wiseness techniques (the criteria adopted to classify the procedures can be found on page55)
6. Flow-charts for each test method was constructed to show how the micro-skills and test-wiseness techniques were combined to form strategies.
7. The two test methods were compared to identify similarities and differences in the micro-skills and test-wiseness techniques.
8. The two test methods were also compared to identify similarities and differences in strategies.
9. Reading processes widely accepted by theoreticians were compared with strategies used by test-takers in order to answer multiple-choice tests to identify similarities and differences.
10. The same reading processes were then compared with strategies used by test-takers in order to answer cloze tests to identify similarities and differences.
11. Based on steps 9 and 10, the construct validity of the two test methods was evaluated.

DISCUSSION OF THE RESULTS

The results of this study will be presented as follows: for both tests, multiple-choice and cloze, first the frequency distribution of correct and incorrect answers and the general index of correct and incorrect answers will be presented. Secondly, the reading micro-skills and test-wiseness techniques which were used by the test-takers in order to answer the two tests will be identified and explained. Thirdly, for each type of test, a flow-chart will be shown of strategies the informants reported that they used in order to come to a conclusion and answer the questions correctly. Fourthly, both types of tests will be compared to detect the similarities and differences between them in terms of micro-skills, test-wiseness techniques and strategies. Finally, current reading processes widely accepted by theoreticians will be compared with strategies used by test-takers in order to answer multiple-choice and cloze tests to verify similarities and differences among them and thereby evaluate these tests' degree of construct validity.

MULTIPLE-CHOICE TESTS

1.Frequency distribution of correct and incorrect answers achieved by informants and their average index of correct and incorrect answers

As mentioned above, there were three multiple-choice tests. Each test consisted of a text and five questions with four alternatives for each question (see appendix 2 pg.101). There were thus 15 questions to be answered, and since there were 23 informants, a total of 345 answers to be analyzed.

Table 1 shows the frequency scores of the correct and incorrect answers and their respective percentages.

Table A : Informants' general correct and incorrect answer index in multiple-choice tests.

Table B

In the last line of the table, we have the total number of correct and incorrect answers and the general index of correct answers: 66.09 compared with 33.91% wrong answers.

2. A flow-chart of the strategies used by the test-takers in order to arrive at the correct answers in the multiple-choice tests.

In chapter one, we saw that, in order to comprehend a text while reading, the reader has to make use of micro-skills in a sequence to form a strategy, and that depending on the strategy a reader uses he will comprehend the text successfully or not.

In this chapter, we will show how the micro-skills and test-wiseness techniques were combined to form strategies.

FLOW CHART 1

1. Micro-skills and test-wiseness techniques used in multiple-choice tests

Flow-chart 1 shows all the micro-skills and test-wiseness techniques the test-takers used in order to compose their strategies and arrive at the correct answers. These micro-skills and test-wiseness techniques are defined below. We will then explain how they were combined to form strategies.

a) **SKIMMING:** we are skimming when we run our eyes over the text, searching for the main points of the text, i.e. its general idea. Skimming, which is a micro-skill, is commonly used while taking a multiple-choice test because the test-taker has to have an idea of what the text is about before actually searching for specific information which may help him answer the questions.

b) **READING STEMS/ALTERNATIVES BEFORE SKIMMING:** while taking a multiple-choice test, some test-takers prefer to read all the stems together with their alternatives before actually reading the text. They probably adopted this test-wiseness technique in order to start reading the text with a specific purpose: trying to identify in the text where probable answers to those items can be found. This technique saves time since the test-taker knows beforehand what he is reading for.

c) **READING STEMS/ALTERNATIVES AFTER SKIMMING:** after having skimmed the text, some test-takers adopt this test-wiseness technique: they generally go back to a stem and its alternatives so that they are able to either answer it straightaway or get more information to narrow down the possibilities. They then return to the text and look for specific information which may lead them to find the correct alternative.

d) **REGRESSION:** this micro-skill seems to serve a different purpose from the original term used by Frank Smith (Chapter 1 pg.7). When taking a multiple-choice test, the test-taker sometimes

needs to go back to parts of the text in order to look for specific information which can lead him to find enough cues in the text to choose the correct alternative.

e) **ANSWER WITH HELP OF PREVIOUS QUESTIONS:** this test-wiseness technique consists of finding enough cues to answer a question on the basis of other questions which have already been answered.

f) **ANSWER STRAIGHT FROM SKIMMING:** some informants skimmed the text once just to get the gist of it and were able to remember enough to choose the correct alternative right after 'reading the question/alternatives after skimming'. Of course this test-wiseness technique depends upon the nature of the question. If the item requires specific/detailed information, the informants most probably will have regressed and scanned to look for it instead of '*answering straight from skimming*'. However if the item requires knowledge of the general idea of the text, then the test-taker may use this micro-skill in order to answer it. This does not mean that only items related to the general idea of the test are answered straightaway. Questions about details of the text might also be answered straight from skimming if the test-taker is able to remember specific parts of the text.

g) **UTILIZE RELEVANT INFORMATION FROM SUBSEQUENT TEST ITEMS AND ALTERNATIVES:** this test-wiseness technique consists of searching for cues in subsequent stems/alternatives in order to find specific information to help them reach the correct alternative.

h) **ELIMINATE ALL THE IMPROBABLE ALTERNATIVES UNTIL ONE IS LEFT WITH THE CORRECT ANSWER:** this test-wiseness technique consists of going back to the text several times to find cues which can help eliminate alternatives which are not mentioned in the text or which carry a different meaning from it. Sometimes the test-takers are able to eliminate all the alternatives which are distractors and therefore are left with only the correct one.

i) **ELIMINATE ALTERNATIVES WHICH ARE KNOWN TO BE INCORRECT AND CHOOSE FROM AMONG THE REMAINING ONES:** when we refer to this test-wiseness technique we mean that the test-taker regressed and scanned (see 'd' p. 61 and 'j' p.62) to look for cues in order to reject one or more alternatives; for instance, alternatives which are not mentioned in the text, or those which contain the opposite idea. Therefore, this technique demands careful search and analysis of the cues in the text. Sometimes the test-taker is left with only one possible relevant alternative. Sometimes he is left with two possibilities, which reduces his chances of making the wrong choice since he is left with a 50% chance (out of two alternatives) if he has to guess the answer. Of course some informants were able to eliminate alternatives without having to regress to the text because they were able to recall the parts they needed to reject or confirm the alternatives they were working on.

j) **SCANNING TO LOCATE SPECIFICALLY REQUIRED INFORMATION:** test-takers use this micro-skill when they read with the aim of extracting only specific information from the text. For instance, it is commonly used when we are looking for a number in a telephone directory. Scanning works together with regression in the same proportion. A test-taker regresses in order to scan and look for cues which can help him reject one or more of the alternatives he is dealing with, or to concentrate on the alternative he suspects is the correct one.

k) **REGRESS AND SCAN TO CONFIRM ANSWER:** some informants decided to regress to the text and scan to look for the cues which confirmed or rejected the alternative they had already chosen as the correct one. This test-wiseness technique was used by cautious test-takers.

l) DEDUCE ANSWER FROM CONTEXTUAL CUES: this test-wiseness technique was used when the informants needed to either find the answer explicitly stated or infer it from the context. Naturally, the sort of contextual cues used in order to reach the correct answer depended on the nature of the items themselves. For example, (item # 10) required the informant to identify the adverb of time and therefore was classified in this research under “lexical item”.

m) INFER FROM READER’S BACKGROUND KNOWLEDGE: this micro-skill was not widely used by the test-takers. In item # 10 , for example, one informant stated that he knew a lot about the subject, as a result of which he was able to answer it without having to go back to the text to look for more information.

n) GUESSING: this test-wiseness technique consists of reaching the correct answer by guesswork, i.e. choosing an alternative without being sure of its correctness. If there are four alternatives a testee will have a one-in-four chance of reaching the right answer by guessing. This technique casts doubts on the validity of this kind of test.

Strategies used in the multiple-choice tests.

Flow-chart 1 also shows how micro-skills and test-wiseness techniques were combined to compose different kinds of strategies which test-takers used in order to achieve the correct answers in the multiple-choice tests.

In the great majority of the answers given, 99.5%, the informants started off by skimming the text, whereas in very few of the cases, less than 1%, they decided to read the stems and alternatives before skimming which was obviously an exception. In the latter case this minority continued their strategies by then *skimming the text* (>1%) followed by the *elimination of alternatives* (>1%), *regressing* again to certain parts of the text in order to *scan* and try to locate information which could help them *deduce the answer from contextual cues* (>1%). In the large number of cases where the informants *skimmed the text* first, the informants continued by adopting the test-wiseness technique *reading the question/alternatives after skimming*. From this step on, six different micro-skills and test-wiseness techniques gave rise to the development of six major strategies, which were divided into minor strategies.

The first major strategy, which began with the micro-skill *regression*, was used in 52.6% of the correct answers. After this micro-skill the test-takers continued by then scanning the text (52.6%), to search for specific information in order to either find cues and *eliminate more alternatives* (23.9%), *deduce the answers from contextual cues* (22.2%), or *regress* again (5.7%) to some parts of the text. In the majority of cases those who chose to *eliminate alternatives* (23.9%) continued their strategy by *deducing the answers from contextual cues* (19.1%). Some of them decided to take a longer path instead, *regressing* (3.5%) to parts of the text in order to *scan* (3.5%) and find specific information to help them *eliminate alternatives* (3%) so they could, from then on, either *deduce the answers from contextual cues* (2.6%) or *regress* for the third time (>1%). The testees who *regressed* continued by *scanning* (>1%) to locate cues and *eliminate alternatives* so they could *deduce the answers from contextual cues* (>1%).

In cases where the test-takers deduced the answers from contextual cues (22.2%), they

were able to find some information while scanning in order to give a definite answer.

The cases in which the test-takers opted to regress again were quite few (5.7%). In such cases they continued by scanning the text in order to search for information so that they could take three different paths. In the first, informants decided to regress (3.5%) again to some parts of the text to scan (3.5%) and from then on either deduce the answers from contextual cues (1.7%) or eliminate alternatives (1.3%). Those who eliminated alternatives ended the strategy by deducing the answers from contextual cues (1.3%). In the other two cases, which represent an exception, the informants either *inferred the answers from their knowledge of the world* (>1%) or *deduced the answers from contextual cues* (>1%).

The second major strategy, which began with the micro-skill reading the stems/alternatives, was adopted by testees who decided to answer the question straight from skimming (27.8%). Of this percentage, in only a very few of the answers did the informants (2.6%) decide to *regress* to the text in order to *scan* and *find information to confirm their answers*.

The third major strategy started off with the test-wiseness technique elimination of alternatives (16.9%). From this point on, in 9.1% of the answers, the informants were ready to answer straight from skimming, while in 6.9% of the answers the test-takers had to regress to the text to scan and from then on either deduce the answer from contextual cues (6%) or regress again to scan (1.3%) and look for information. They continued by either eliminating alternatives (>1%), so they could deduce the answer from contextual cues (>1%) or regressing (>1%) and scanning (>1%) for the third time to find information which enabled them to deduce the answers from contextual cues. (>1%)

The fourth, fifth and sixth major strategies were followed by testees who decided to continue their paths after having read the stems/alternatives. Very few informants used these strategies to reach their answers: less than 1% of the answers given. Some informants answered the question with the help of previous questions. Others decided to eliminate improbable alternatives until they were left with only one correct alternative. The third group *utilized relevant information from subsequent test items* followed by *regression* and *scanning*, *eliminating alternatives* and *deducing answers from contextual cues*.

THE CLOZE TEST

Frequency distribution of correct and incorrect answers achieved by informants and the average index of the correct and incorrect answers

As explained in chapter three, a cloze test was selected from 'The University of Cambridge International Examinations'- Certificate of Proficiency in English - June 16, 1993 (see appendix 3 p.108). We changed the deletion rate from a rational deletion criterion (see Chapter 2 pg 45) to a fixed-ratio deletion (see Chapter 2 pg.45), i.e. we mechanically deleted every 6th word. Twenty words were deleted. We left the first and last two sentences intact so the reader could have a better idea of what the content was about. Since there were 23 informants and 20 blanks to be filled in, we had a total of 460 items to be analyzed. Again only the correct answers

were analyzed because we were only interested in micro-skills and strategies which led to the correct answers.

Table 2 shows the frequency scores and the average index of the correct and incorrect answers the test-takers achieved.

Table 2: Informants' general correct and incorrect answer index in the cloze test.

TABLE 2

FLOW CHART 2

Micro-skills and test-wiseness techniques used in the cloze test and their explanation.

- a) **SCANNING:** The testees claimed to have used this micro-skill (see 'j' p.62) in order to locate detailed information which they needed to complete the blank.
- b) **SKIM A PARAGRAPH:** skimming here continues being the search for the central idea of the text. However, it is possible to look for the main ideas of each paragraph. Some informants decided to use this micro-skill in order to have a general idea of what that paragraph was about before working further on it.
- c) **SKIM THE WHOLE TEXT:** (see 'a' p. 60)
- d) **USE OF LANGUAGE KNOWLEDGE:** This micro-skill was definitely the most commonly used. The test-takers had to work with this skill not only in order to analyze the blanks, but also the sentences and paragraphs. These analyses led them to reach the correct answers.
- Some informants claimed to have used only knowledge of grammar such as verb tenses or preposition use. However, when one is trying to fill in blanks in cloze tests, it seems that the test-taker is using different kinds of grammatical knowledge at the same time, even if he is not aware of that. It seems that not only does a test-taker use syntactic analysis, but also semantic analysis (not only the word in relation to the text, but also collocation, i.e.the way that some words occur regularly whenever another word is used.).
- e) **REGRESSION:** (see 'd' p. 61)
- f) **READER'S PREVIOUS KNOWLEDGE OF THE SUBJECT:** Some test-takers claimed to have used this micro-skill which helped them infer the answers to some items. For instance, blank# 04 was about the Greek and Latin languages. One informant, who is a history teacher, wrote that when she read `_____ and Latin' , Greek came to her mind immediately, because in history these two cultures were interrelated and always came together in history books. Other informants who did not have the same background were misled into supplying the word 'English' just because they were taking a reading comprehension test in the English language.
- g) **SKIMMING FORWARD:**Some informants had to stop trying to fill in the blank where they were and skim forward, i.e. read for a general idea of the following part of the text, in order to

understand the paragraph so that they could regress to the blank they were stuck on and attempt to answer it. Although skimming is a micro-skill related to reading, skimming forward seems to be a test-wiseness technique because testees use it in order to help them fill in the blank(s).

h) GUESSING: (see 'n' p. 63)

i) UTILIZE RELEVANT INFORMATION FROM PREVIOUS TEST ITEMS:

j) ANSWER ITEMS AHEAD FIRST AND THEN REGRESS TO ANSWER BLANK IN QUESTION: In this case only one or two blanks, which were ahead of the one the testees were stuck on, were answered. Right after that the test-taker regressed immediately and tried to link what has been answered to the blank he was working on. This test-wiseness technique sometimes helped the testee to get a better idea of what the text was about.

k) REGRESS AND SCAN TO CONFIRM ANSWER: see ('k' p. 63)

l) READ PREVIOUS AND LATER SENTENCES: This test-wiseness technique consists of regressing to a previous sentence and reading it and then from where you are (i.e. the blank you are at) going forward to read the next sentence before going back to the original blank and completing it.

m) AFTER HAVING FINISHED THE TEST GO BACK TO COMPLETE BLANKS WHICH WERE LEFT UNANSWERED: some test-takers were not able to fill in some blanks. Therefore, they left these blanks aside and continued completing the remaining ones. When they finished the test they went back to the blanks which were left incomplete and tried to complete them. This test-wiseness technique helped them because once most of the text was filled in they were able to have a very good idea of what the text was about, which definitely helped them choose an appropriate word for those blanks which had been left incomplete.

n) INDEFINITE ANSWERS: some test-takers were not able to find appropriate words for some blanks and left them unanswered. Others were not able to describe how they found the correct answers. Perhaps the cloze test is one of the most difficult tests to understand in terms of what it really measures and how one comes to a plausible answer.

Strategies used in the cloze test.

Flow-chart 2 shows how the informants arrived at the correct answers by using different strategies.

In 83.47% of the answers given, the informants started their tests by scanning in order to read for details, as against a minority who took two different paths. In 3.36% of the cases, the informants started their tests by skimming one paragraph. In only 1.12% of the answers given the test-takers started by skimming the whole text.

In the majority of the cases, the informants who began by scanning to read for details, continued their strategies by taking six different paths:

First, 59.38% of the correct answers were achieved by informants who used language knowledge to complete the blanks. Few of the the answers (1.12%) were reached by informants who decided to regress and scan to confirm the answers.

Secondly, in a minority of the cases (4.48%), the informants regressed to parts of the text and from there on took four other paths: 1.96% of the answers were achieved by informants who *scanned* to locate specific information which could help them, together with the *use of language knowledge* (1.4%), to complete the blanks. In less than 1% of the cases, the informants decided to *regress* again and *scan* to locate specific information and *use their knowledge of the language* to come to an answer. In other answers (1.12%) the informants *skimmed the paragraph, regressed* (1.12%) to parts of the text to either *use knowledge of the language* (>1%) to come to an answer or *skim the paragraph* (>1%) again. From then on less than 1% of the answers were reached by *regressing* for the third time. Then they decided to *skim forward* (>1%) in order to find clues so they could *use their knowledge of the language* (>1%) and complete the blank. In a small minority of the answers given, less than 1%, informants *scanned previous and later sentences* so they could find cues and *use knowledge of the language* to arrive at an answer. In other cases still, (>1%), the informants decided to *skim forward* and then *regress* in order to *scan* and locate cues to *use knowledge of language* in order to complete their strategy.

Thirdly, in a minority of the answers given - 2.8% - the informants answered the blanks based on their knowledge of the world.

Fourthly, in 2.52% of the cases, the informants skimmed forward, regressed (2.52%) to either scan (>1%) to locate specific cues so they could use their knowledge of the language (>1%) to answer the items or to skim a paragraph (>1%) and from then on use knowledge of the language (>1%) to answer the item adequately.

Fifthly, in 1.96% of the answers given, the informants claimed to have guessed the correct answers.

And finally, in less than 1% of the answers given, the informants claimed to have utilized relevant information from previous items to find cues which could lead them to use their knowledge of the language and answer the items correctly.

The testees who decided to start answering the test by skimming a paragraph (3.36%) took two other paths. Some (2.8%) regressed in order to scan (2.52 %) and locate information which could lead them to use knowledge of language (2.52%) and complete the blanks. Others *skimmed a paragraph* (>1%) to be able to make *use of language knowledge* (>1%) and answer the test. In less than 1% of the cases, the informants *answered items ahead*, followed by *regression and scanning*, and finally made *use of their language knowledge* to complete the blanks.

The third group's answers were achieved by testees who started their tests by skimming the whole text (1.12%). They continued by regressing (1.12%), scanning (1.12%) and using knowledge of the language (1.12%) to reach the correct answers.

Again, if we view the flow-chart as a whole we may conclude that since there are various different strategies a test-taker can adopt in order to reach a correct answer, the percentage of frequency of these strategies tends to decrease owing to the various paths one can take.

COMPARISON OF MULTIPLE-CHOICE WITH CLOZE TESTS

Multiple-choice and cloze procedures lead test-takers to use strategies which are peculiar to each test type. In multiple-choice tests subjects tended to use strategies which included micro-skills and test-wiseness techniques such as skimming the whole text (99.5%), and reading alternatives after skimming (99.5%). We may see that in this type of test there are three main typical branches a test-taker follows after reading alternatives after skimming: regression (52.6%), answering straight from skimming (27.8%) and elimination of alternatives (16.9%). In the cloze test, subjects tended to start their strategies using micro-skills such as *scanning to read for details* (83.47%) and *use of language knowledge* (59.38%). The comparison of the two flow-charts shows that the combination of the micro-skills and test-wiseness techniques into strategies is peculiar to each kind of test. Each type of test led test-takers to use particular strategies because of the format of these tests, that is, the way and order in which they are arranged and presented. Therefore, there are no strategies which are common to both tests. That is to say, the strategies used to answer multiple-choice or cloze tests consist of either different ordering of micro-skills and test-wiseness techniques which are common to both tests, or different ordering of these skills and techniques which is specific to each test. Consequently, the strategies these test-takers used in order to reach the correct answers were peculiar to each type of test. Below are the micro-skills and test-wiseness techniques which were used to compose the strategies of the two tests. We will divide this section of the discussion of the results into three parts:

- i) micro-skills and test-wiseness techniques common to both tests;
- ii) micro-skills and test-wiseness techniques peculiar to multiple-choice tests;
- iii) micro-skills and test-wiseness techniques peculiar to cloze tests.

When comparing the two flow-charts, one can perceive that some test-takers used the same micro-skills and test-wiseness techniques in both tests while others were used exclusively in either the multiple-choice or cloze tests.

The micro-skills and test-wiseness techniques which were used in both tests

a) regress and scan for specific information: when the questions or blanks required specific information to answer them, testees were forced to return to parts of the text in order to scan it and try to locate the necessary information. These micro-skills were commonly typical to both tests. 52.6% of the answers given in multiple-choice tests contained this micro-skill and 83.47% in the cloze test. In both tests there was significant use of *regression* and *scanning for specific information*, because both tests cause testees to search for more information in the text they are reading if they intend to find the correct answers.

b) skim text: this micro-skill was used when the testees wanted to have an overall view of what the text was about. Once they had a general idea of the whole text it made it easier to work on parts of it, i.e. look for details. Although this micro-skill was common to both tests it was not typical of both of them. 99.5% of the answers given in the multiple-choice tests contained this

micro-skill, whereas only 1.12% of the answers given in the cloze test contained it.

c) infer the answer from knowledge of the world: when the testee had previous knowledge of the subject of the text, it was easier at times to predict the answers on the basis of this existing knowledge. Although this micro-skill was common to both tests it was not typical. This micro-skill was used in less than 1% of the answers given in the multiple-choice tests and in the cloze test only 2.8% of the answers were achieved by utilizing this micro-skill.

d) regress and scan to locate specific information to confirm answer : some testees found it safer to return to the text again to scan and locate information where they could confirm their answers. Although this test-wiseness technique was common to both tests, it was not typical. This technique was used in 2.6% of the answers given in the multiple-choice tests and 1.12% in the cloze test.

5.3.2. Micro-skills and test-wiseness techniques which were restricted to the multiple-choice tests

a) read stems/alternatives after skimming : the format of the multiple-choice test determines the use of this test-wiseness technique. This test-wiseness technique was typical: it was used in 99.5% of the answers given.

b) eliminate alternatives : this technique was used in many of the strategies and sometimes more than once in one strategy. Together with *regression*, *scanning* and *deducing answer from contextual cues*, *eliminating alternatives* was one of the most common test-wiseness techniques used because test-takers found it safer to *eliminate alternatives*. They were left with fewer options to choose from, increasing the probability of getting the answer right. For instance, if a testee is left with two alternatives, even if he has to guess which one is correct, he will have a 50% chance of guessing the correct alternative.

c) answer straight from skimming : usually when general comprehension questions were asked testees were able to retain enough information from the first reading to answer the question soon after.

d) eliminate improbable alternatives until you are left with only one correct alternative: Although this technique is very similar to 'eliminating alternatives', in the sense that test-takers eliminate options by analyzing parts of the text, it differs because, here, the testee is able to discount all distractors and reach the correct answer by a process of elimination.

e) answer with help of previous question : answers to previous questions allowed test-takers to have a better comprehension of the text, which facilitated their answering of other questions.

f) deduce answer from contextual cues : contextual cues here made it possible for test-takers to find out more about the message. Sometimes an adverb of time helped them answer the questions. At other times a small part of the text contained essential information.

g) read questions/alternatives before skimming : this test-wiseness technique is only used when there are alternatives to choose from, which is not the case in cloze tests.

These test-wiseness techniques are peculiar to multiple-choice tests because they are a reflection of the format of this kind of test. Some of them are typical, that is, they happen more often than others. If there are questions and alternatives, for instance, the test-taker has to either read them before reading the text or after reading it - most of the informants (99.5%) preferred to

read them after skimming the text. Another typical technique, which occurred once or twice in various strategies, was elimination of alternatives. Test-wiseness techniques presented here were therefore only used in the multiple-choice tests because of the particular characteristics of this type of test.

Micro-skills and test-wiseness techniques which were confined to the cloze test

a) use of knowledge of the language : most of the informants claimed to have used syntactic and semantic analysis, verb tense knowledge, and other kinds of grammatical knowledge in order to be able to complete the blanks. This micro-skill was peculiar to the cloze test owing to its format, i.e. when filling in blanks one must have a notion of the structure of the language to find out what is missing and why. It was also the skill which was most used by the informants to reach the correct answers. All the strategies (100%) ended with this skill showing its supremacy in this type of test.

b) answer items ahead: some informants found it necessary to complete the text which was ahead of the items they were working on so they could have a better idea of the message of the text. Of course one can leave one question behind and answer the others first in the multiple-choice test; however, this did not happen with our informants in this kind of test, but only with the cloze test. Perhaps it happens more often in cloze tests owing to the importance of having to have the whole filled in to be able to understand its parts.

c) skim forward : also peculiar to the cloze test, because when a testee gets stuck on a blank he feels the need to quickly read ahead in order to comprehend the text in depth and maybe be able to go back and answer the blank he had problems with.

d) scan previous and later sentences : since cloze tests cause test-takers to work more at sentence level, they often read the sentences which precede and follow the one they are working on in order to look for details which might help them complete the blank .

Again, the micro-skills and test-wiseness techniques which were exclusively used in the cloze test were determined by the characteristics of this kind of test. Some techniques and skills were not typical, such as answering items ahead, skimming forward, scanning previous and later sentences; however, use of language knowledge occurred in all strategies.

COMPARISON OF CURRENT READING PROCESS THEORIES WIDELY ACCEPTED BY THEORETICIANS WITH STRATEGIES USED IN MULTIPLE-CHOICE AND CLOZE TESTS

In this section, current reading process theories widely accepted by theoreticians will be compared with the strategies used in multiple-choice and cloze tests which were identified in this research. Firstly, we will show figures illustrating different theoreticians' views of the reading process. Secondly, we will compare these figures to the multiple-choice test strategies. Thirdly, we will compare the same figures to the cloze test strategies. Finally we will compare the strategies used in the multiple-choice and cloze tests with normal reading strategies.

When comparing the reading process described by different theoreticians, whose views are summarized in the figures above, to the test-taking strategies, we are able to perceive that both multiple-choice and cloze reading comprehension tests demand a tremendous effort from the

testees: much more than if they were merely reading the text. The figure below shows the paths readers take when simply reading or when taking multiple-choice and cloze reading comprehension tests.

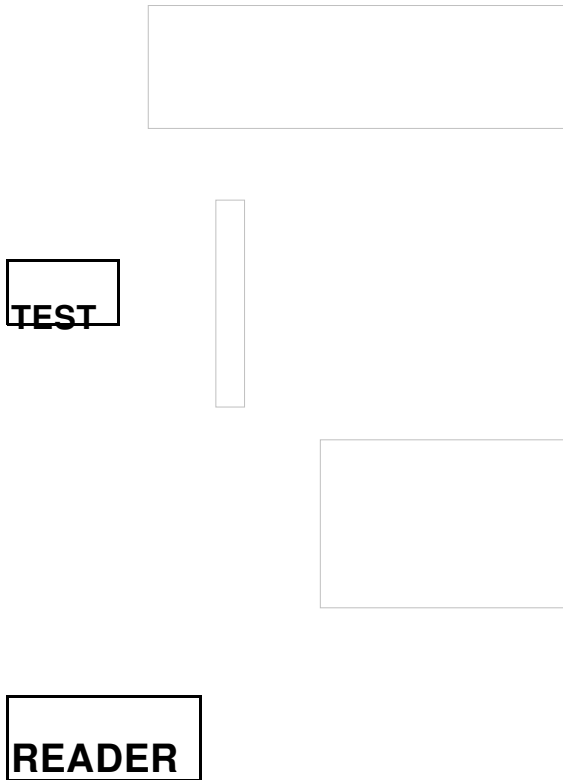


Figure 15: The paths readers take when simply reading () and when taking a multiple-choice or cloze reading comprehension test ().

We can notice that when testees take a multiple-choice or cloze test, the path they take in order to interpret the message is different from a common interactive approach to reading. In the flow-chart of the cloze test, we can perceive that all the strategies end with the micro-skill 'knowledge of the language'. This indicates that the test-takers concentrate their reading on a 'visual' (bottom-up) level. Therefore, if we take a look at figure 11 we will see that the testee utilizes more 'visual' than 'non-visual information'. This causes overloading of short-term memory (see p.14) and consequently a 'bottleneck effect' (see p.9) and then 'tunnel vision' (see p.10). In figure 12 we can see that when a testee is taking a cloze test he spends more time on the 'optical', 'perceptual' and 'syntactic' cycle, neglecting 'meaning'. This means that more time is spent on the attempt to fill in the blanks than on actually comprehending the message of the text as a whole. According to figure 13, people-taking a cloze test devote too much time to 'syntactic' and 'semantic knowledge', making the 'pattern synthesizer' unstable in the sense that instead of quickly analyzing the features in four ways (syntactic, semantic, orthographic and lexical) they spend too much time on two of them: 'syntactic and semantic'. In figure 14,

test-takers of a cloze test work almost exclusively on the knowledge of the language (bottom-up process), interrupting the interaction with the knowledge of the subject and world (top-down process).

Multiple-choice tests also present problems. If we compare figures 11, 12, 13 and 14 to flow chart 1 we will notice that besides getting engaged in an interactive form of reading, multiple-choice test-takers also get trapped in the arduous task of choosing a correct alternative to a question. Some items make testees regress to the text and read it in an interactive way. For instance, item #1, in the first passage, causes testees to return to the text and re-read paragraph one in such a way that they are able to apprehend the general meaning of that part. Other items induce testees to carry out a bottom-up analysis of the text. For instance, in item #10, in passage two, the testee inevitably has to find and understand the word 'nowadays' in order to arrive at the correct alternative. Sometimes, the answers are found in one single sentence as in item #14, in passage three, ('America sells the steak, while Britain sells the sizzle'). We will have an interactive form of reading in multiple-choice tests, depending on the nature of each question. Some questions will lead to interactive reading approach while others will narrow the process down to the bottom-up level as described in connection with the cloze tests. However, even though a test writer spends time and effort pre-testing and analyzing the items, multiple-choice tests continue leading testees to use test-wiseness strategies which may lower the validity of this kind of test. Weir (1990), Hughes (1994) and Oller (1986) also show other problems related to multiple-choice tests. They explain why multiple-choice tests are time-consuming and impractical. These theorists also claim that a test-writer will face problems such as writing subjective questions, finding suitable distractors, and so on.

We can therefore posit that both multiple-choice and cloze tests have low construct validity. Cloze tests induce test-takers to work on their reading comprehension in a bottom-up way. While taking a cloze test, testees usually get so bogged down in the task of analyzing the structures of the sentences in order to find the appropriate missing word that they lose track of the message of the text, which is not the objective when one is attempting to measure reading comprehension. This phenomenon, i.e. the cloze test inducing test-takers to use bottom-up strategies, is understandable if we consider when it was devised. W.L.Taylor introduced the cloze procedure in 1953, at that time when reading was first analyzed into its basic parts, which were then taught separately. For instance, readers had to identify individual letters, building them up to words, sentences and then paragraphs. The cloze procedure reflected the philosophy of that period. It makes a testee work at word level rather than at text level. Multiple-choice tests have low construct validity because the test-takers have to concentrate on composing strategies which enable them to select the correct alternative rather than on normal reading strategies.

In short, choosing an appropriate test type and preparing the test is a demanding job. A test writer should be aware of all the complexity which tests involve before embarking on the task.

CONCLUSION

The objective of this dissertation was to first of all make an inventory of strategies testees used while taking multiple-choice and cloze tests, and check if these types of tests induced test-

takers to use particular strategies. We then compared current reading process theories with multiple-choice and cloze test strategies, to establish which strategies were common to both tests and which were specific to one or the other type.

Firstly, it was found that multiple-choice and cloze tests each cause test-takers to use particular strategies in order to reach the correct answers, although some strategies are composed of micro-skills and test-wisness techniques which are common to both types of test.

Secondly, both multiple-choice and cloze tests demand more than reading abilities: they cause test-takers to use strategies which are specific to each type of test. Students taking multiple-choice and cloze tests not only become engaged, even if partially, in the reading process as described by Smith (1978a, 1978b), Goodman (1976), Eskey (1989), Rumelhart (1977) and Stanovich (1980), but also in a test-taking process which demands the use of particular test-taking strategies (specific to each test) in order to reach their objective of arriving at the correct answers.

Finally, construct validity might be reduced in multiple-choice and cloze tests since they induce test-takers to use test-wisness techniques, which is not what these tests are intended to measure.

BIBLIOGRAPHY

ALDERSON, J. Charles; HUGHES, Arthur (Eds.) **ELT documents III - Issues in language testing.**

London : The British Council, 1981.

ALDERSON, J. Charles; URQUHART, A.H. **Reading in a foreign language.** New York : Longman, 1984.

ALDERSON, J. Charles; CLAPHAM, Caroline, WALL, Dianne. **Language test construction and**

evaluation. Cambridge : Cambridge University Press, 1995

ALDERSON, J. Charles. The cloze procedure and proficiency in English as a foreign language. **TESOL**

Quarterly. Mexico City, n.13, p.219-26, 1979.

_____. Native and nonnative speaker performance on cloze tests. **Language Learning.** Lancaster,

v.30, n.1., p.59-76, 1980.

_____. **Testing reading comprehension.** Lancaster, Centre for Research in Language Education -

Working Paper 1, 1992.

_____. **Judgements in language testing.** Lancaster, Centre for Research in Language Education-

Working Paper 3, 1992.

_____. **The relationship between grammar and reading in an EAP test battery.** Lancaster, Centre

for Research in Language Education-Working Paper 2, 1992.

ALLAN, Alastair. Development and validation of a scale to measure test-wisness in EFL/ESL reading

test takers. **Language Testing**. Hong Kong, v.9, n.2., 1992. p.101-122

ANDERSON, Richard C.; PEARSON, P. David. A schema-theoretic view of basic processes in reading

comprehension. In: CARRELL, Patricia L.; DEVINE, Joanne; ESKEY, David E. (Eds)

Interactive

Approaches to second language reading. Cambridge : CUP, 1988.

BARROS, Anna Rachel Machado Paes de. Estrategicamente falando de estratégias....

Trabalhos em

Lingüística Aplicada, Campinas, n.5-6, p.25-39, jun./dez. 1985.

BACHMAN, Lyle F. What does language testing have to offer? **Tesol Quarterly**. California, v.25, n.4,

Winter 1991. p.671-704

_____. Performance on cloze tests with fixed-ratio and rational deletions. **Tesol Quarterly**. Illinois,

v.19, n.3, September, 1985. p535-556

BASTIDAS, Jesus A. The cloze procedure as a teaching technique to improve reading comprehension.

English Teaching Forum, Narino, Colombia, v. 22, n.2, p.20-31, April 1984.

BAKER, David. **Language testing - A critical survey and practical guide**. London : Edward

Arnold, 1989.

BRAGA, Denise B.; BUSNARDO, Joanne. Metacognition and foreign language reading: fostering

awareness of linguistic form and cognitive process in the teaching of language through text.

Lenguas Modernas. Chile : Universidad de Chile, n.20, 1993, p.129-49.

CARRELL, P. Three components of background knowledge in reading. **Language Learning**. Illinois,

v.33, n.2, 1983. p.183-207.

_____. Some causes of text-boundedness and schema interference in ESL reading. In: CARRELL, Patricia

L., DEVINE, Joanne; ESKEY, David E. (Eds). 1988.

CHAVEZ-OLLER, M.A.; et al. When are cloze items sensitive to constraints across sentences?

Language Learning. City of Albuquerque, v.35, n.2, 1985.

COHEN, Andrew D. **Testing language ability in the classroom**. Massachusetts : Newbury House

Publishers, 1988.

_____. On taking tests: what the students report. **Language testing**. n.1, 1:70-81, 1984.

DAVIES,

Alan (Ed.) **Language testing symposium: a psycho-linguistic approach**. London : Oxford University Press, 1968

_____. Reaction to the Palmer & Bachman and the Vollmer papers (2). In: ALDERSON, J.C.; HUGHES, A. (Ed). **ELT Documents III** - Issues in language testing. London :The British Council,

1981. p.182-186

DOUGHTY, C; PICA, T. Information gap tasks: do they facilitate second language acquisition?.

TESOL Quarterly, University of Pennsylvania, v.20, n.2, p.305-24, 1986.

- ECO, U. **The role of the reader: explorations in the semiotics of texts.** Bloomington : Indiana University Press, 1981.
- ELLIS, Rod. **The study of second language acquisition.** Oxford : Oxford University Press, 1994.
- _____. **Understanding second language acquisition.** Oxford : Oxford University Press, 1985.
- ESKEY, D. Holding in the bottom. In: CARRELL, Patricia L.; DEVINE, Joanne, ESKEY, David E. (Eds). 1988. _
- GOODMAN, Kenneth S. The psycholinguistic nature of the reading process. In: GOODMAN, Kenneth S. **The psycholinguistic nature of the reading process.** Detroit : Wayne State University Press, 1973. p.13-26
- _____. Behind the eye: what happens in reading. In: SINGER, Harry; RUDDELL, Robert B (Eds).
- Theoretical models and processes of reading.** Newark, Delaware : International Reading Association, 1976. p.470-496
- _____. Reading: a psycholinguistic guessing game. In: SINGER, Harry; RUDDELLI, Robert B. (Eds.).
- Theoretical models and processes of reading.** Newark, Delaware : International Reading Association, 1976. p.497-508
- _____. The reading process. In: CARRELL. Patricia L.; DEVINE, Joanne; ESKEY, David E. (Eds). 1988.
- GRIGOLETTO, Marisa. A influência da previsibilidade sobre as estratégias de compreensão na leitura em língua materna e língua estrangeira. **Trabalhos em Lingüística Aplicada,** Campinas, n.15, p.25-47, jan./jun. 1990.
- HARRIS, David P. Testing reading comprehension. In: **Testing English as a second language.** New York : McGraw-Hill Book Company, 1969.
- HEATON, Brian. **Classroom testing.** London : Longman, 1994.
- _____. **Writing English language test.** London : Longman, 1975.
- HENNING, Grant. Dimensionality and construct validity of language tests. **Language Testing.** London, v.9, n.1, 1992.
- _____. **A guide to language testing development, evaluation, research.** London : Heinle & Heinle, 1994.
- HUGHES, Arthur. **Testing for language teachers.** Cambridge: Cambridge University Press, 1994.
- JONZ, John. Another turn in the conversation: What does cloze measure? **TESOL Quarterly.** Texas, v.24, n.1, Spring 1990.

LYLE, Bachman. **Fundamental considerations in language testing**. Oxford : Oxford University

Press, 1994.

MADSEN, Harold S. **Techniques in testing**. New York : Oxford University Press, 1983.

MEURER, José Luiz. Efeitos dos organizadores antecipatórios na leitura em língua estrangeira e

língua materna. **Trabalhos em Lingüística Aplicada**, Campinas, n.10, p.9-36, jun./dez. 1987.

OLLER JR., John; PERKINS, Kyle. **Language in education: testing the tests**.

Massachusetts :

Newbury House Publishers, 1978.

_____. **Language tests in school**. London : Longman, 1986.

_____. **Issues in language testing research**. Massachusetts : Newbury House Publishers, 1983.

RUMELHART, D. Toward an interactive model of reading. In: KORNBLUM, Sylvan (Ed).

Attention

and Performance. vol. VI, New York: Academic Press. 1977. p.573-603

SAMUELS, S. Jay; KAMIL L. Models of the reading process. In: CARRELL, Patricia L.; DEVINE,

Joanne; ESKEY, David E 1988.

SCARAMUCCI, Matilde V. R. O resumo e a avaliação da compreensão em leitura em língua estrangeira. **Trabalhos em Lingüística Aplicada**, Campinas, n.15, p.65-86, jan./jun. 1990.

SCHOHAMY, R. Does the testing method make a difference? The case of reading comprehension.

Language Testing, Tel Aviv, v.1, n.2, p.147-70, 1984.

SILBERSTEIN, Sandra. Let's take another look at reading: 25 years of reading instruction.

English

Teaching Forum. Washington D.C., Oct, 1987. p.28-35.

SHOHAM, Míriam; PERETZ, Arna; VORHAUS, Renee. **Reading comprehension tests: general or**

subject-specific? Ben Gurion University of the Negev, Israel. 1985.

SHOHAMY, Elana. **A practical handbook in Language testing for the second language teacher**.

Tel-Aviv, Tel-Aviv University, 1985.

SMITH, Frank. **Reading**. Cambridge : Cambridge University Press, 1978a.

_____. **Understanding Reading: a psycholinguistic analysis of reading and learning to read**. New York : Holt, Rinehart and Winston, 1978b.

SOUDEK, Miluse; SOUDEK, Lev I. Cloze after thirty years: new uses in language teaching. **ELT**

Journal, Oxford University Press , v.37, n.4, October 1983. p335-341.

STANOVICH, K.E. Towards an interactive-compensatory model of individual differences in development of reading fluency. **Reading Research Quarterly**, n.16, 1980. p.32-71

WEIR, Cyril. **Communicative language testing**. New Jersey : Prentice Hall International, 1990.

_____. **Understanding and developing language tests**. New Jersey : Prentice Hall International, 1994.

WIDDOWSON, H. **Teaching Language as Communication**. Oxford : OxfordUniversity Press, 1978.

[1] K. Goodman, 1967

[2] An experienced British EFL professor, Michael Alan Watkins-UFPR, judged which words were acceptable or not in the cloze test.