

THREE FACTORS TO CONSIDER IN THE ASSESSMENT OF EFL LEARNERS' READING COMPREHENSION ABILITY: THE READER, THE TEXT, AND THE TASK

YABANCI DİL ÖĞRENİMİNDE ÖĞRENCİLERİN OKUMA-ANLAMA BECERİLERİNİ ÖLÇMEDE GÖZ ÖNÜNDE BULUNDURULMASI GEREKEN ÜÇ ETKEN: OKUYUCU, OKUMA PARÇASI, VE DE SORU TÜRÜ

Berrin UÇKUN*

ABSTRACT: This study was carried out with the intention of examining the three factors - the reader, the text and the task - which were expected to play a crucial role in the assessment of EFL learner's reading comprehension ability. The data for the study were collected from university students in prep. English programs. The results of the study reveal positive support for using more than one technique per text to measure reading ability. The multiple-choice task, consisting of factual, inferencing, and generalization questions, and the written recall task are examined in their interaction with two text types. An analysis of data obtained from these two task and text types draw attention to the importance of reader's familiarity with a given text topic and how this can affect results on two measurement techniques. Qualitative analysis of written recalls reveal interesting information on the text processing strategies of students.

Key Words : *Written recall protocols, expository texts, content familiarity, assessment tasks.*

ÖZET: Bu çalışma İngilizce yabancı dil eğitimi alan öğrencilerin yabancı dilde okuma-anlama becerilerinin ölçümünde önemli rol oynadığı düşünülen üç etkenin incelenmesi amacı ile yürütülmüştür - metin, okuyucu, ve ölçüm teknikleri. Çalışmayı oluşturan veriler İngilizce hazırlık sınıfı okuyan üniversite öğrencilerinden toplanmıştır. Çalışma sonuçları okuma becerisinin ölçümünde birden fazla yöntem kullanılmasını kuvvetle desteklemektedir, ve tercihen bu yöntemlerin bazıları okuyucunun ne anladığını kendi sözcükleri ile ifade etmesi şeklinde olmalıdır. Kelimesi kelimesine bilgi, anlam çıkarma, ve genelleştirme türünde oluşturulan çoktan seçmeli sorular ve akılda kalanın yazılı ifadesi şeklinde oluşturulan ölçüm yöntemlerinin okuma metnindeki farklılıklar ile etkileşimi incelendi. Ortaya çıkan sonuç, okuyucunun okuma metnindeki bilgilere aşinalığının önemini ve bunun ölçüm sonuçlarını nasıl et-

kilediğini göstermektedir. Yazılı anlatımların nitel çözümlemesi sonucu öğrencilerin metin çözümleme yöntemleri konusunda ilginç bilgiler ortaya konulmuştur.

Anahtar Kelimeler : *Okunan parçadan hatırlanan yazılı anlatım, düz yazım metinleri, içerikle aşinalık, ölçüm yöntemleri*

1. INTRODUCTION

There is an abundance of research related to second and foreign language reading. However, the findings of these research appear to be quite disparate because researchers vary greatly in the comprehension assessment procedures they utilize [1], their subjects' existing knowledge of the texts, and the structural features of the texts they use in their studies. Wolf [2] points to the importance of achieving comparability of findings before we can adequately support or disprove existing theories. Therefore, in order to be able to reach conclusions regarding the findings of our studies and compare results across studies, research must determine how the following three factors affect a foreign language reader's performance on any reading task: the characteristics of the text, the reader's prior knowledge of the text content, and the nature of the assessment task.

For a long time, text analysis was limited to analyses of text difficulty and syntactic complexity, whereby the sentence was used as the unit

* Yrd. Doç. Dr. İngiliz Dili ve Edebiyatı A.B.D. Gaziantep Üniversitesi

of analysis. Text linguistics, on the other hand, developed with the purpose of analyzing texts into units which are longer than single sentences. It provided "a context of awareness for examining texts at a variety of levels and in a variety of modes" [3]. Three of the techniques most often used in reading research are propositional analysis, cohesion analysis, and story grammars. Work on story grammar revealed the existence of conventionalized macrostructures which derive from background knowledge of texts. Unlike narrative texts, the macrostructures in expository texts are more difficult to describe because the relationships among text segments are more diverse due to a greater variety of possible content matter. Carrell [4] searched for the effects of four different English rhetorical patterns on ESL readers' recalls. Her findings confirmed that the more tightly organized patterns of comparison, causation, and problem/solution generally facilitated the recall of specific ideas from a text more than did the collection of description, despite differences in the subjects' language backgrounds. In a later study, Carrell [5] investigated whether there were differences in quantitative and qualitative analysis of reading recall protocols as a function of different text structures (i.e. comparison/contrast versus collection of description). There were no significant differences between the two types of text structures in the quantity of information recalled, but a qualitative analysis revealed that there were significantly more top-level idea units recalled from the comparison/contrast passage.

Schema theory [6] has highlighted the different effects of background knowledge in reading and stressed the need for culturally appropriate instructional reading materials. Carrell [7] argued that "implicit content knowledge presupposed by a text and a reader's own cultural background knowledge of content interact" (p.104) with the result that understanding a text based on one's own culture becomes easier to understand than a text which is syntactically and rhetorically equivalent but based on a less familiar distant culture. Johnston [8] made a distinction

between quantitative and qualitative differences in background knowledge. A quantitative mismatch refers to a lack of relevant world knowledge. When this is the case, the readers cannot make appropriate inferences when they are required, and cannot tie the new information in the text to knowledge structures already existing in the reader's cognitive structure.

Chen and Donin [9] have carried out a study investigating the effects of language proficiency and domain-specific knowledge on discourse processing. The authors have reached the conclusion that in the comprehension of semantic knowledge conceptual knowledge plays a much more important role than linguistic knowledge, and also exerts much stronger compensatory effects than realised before. It is necessary to consider the implications of these studies for reading assessment, especially when interpreting performance on multiple-choice questions, inference questions, and when answering questions with the text not available for referral. Without information on prior knowledge, we could not be sure whether the failure was due to a lack of prior knowledge or failure to use available prior knowledge. Lack of familiarity with text structure, content matter, and structural cues all make the reading tasks more difficult for the reader.

L2 reading research has shown that learner's performance on comprehension task also varies according to the task used [10]. Wolf [2] points out that such variation can be attributed to two factors: the type of task used to assess comprehension and the language in which the task is presented and carried out. In a study where she examined the effects of task type, language of assessment, and target language experience on learners' ability to demonstrate comprehension on post-reading tasks, Wolf found that subjects at both the beginning and advanced levels performed better on selected response tasks than on constructed response tasks. She argues that tasks requiring construction of responses in the target language have a debilitating effect on test-takers' performance. Wolf's assertions were confirmed by the findings of the studies carried out

by Shohamy [11] and Gordon and Hanauer [12]. Lee [10] suggests that when conducting research on the reading comprehension of learners with different levels of target language experience, a combination of tasks yields more information on subjects' comprehension than one task alone. Furthermore, it may be beneficial to test learners in their native language if we want to obtain an accurate estimate of the subject's ability to comprehend target language texts.

One major issue of debate regarding reading assessment is whether reading comprehension should be viewed as a process or a product (see Bernhard [13] & Johnston, [8]). Royer and Cunningham (in Johnston, [8]) state that the product and process approaches are complementary approaches to the reading problem because a comprehended message will be kept in memory better than an uncomprehended message. The processes of reading comprehension can only act upon information already stored in the memory, meaning that process and product must be studied in their interaction.

This study focused on the assessment task factor while paying special attention to a) the readers' existing background knowledge related to the contents of the test passages and b) the comparability of the two texts from syntactic, semantic and rhetorical points of view. Two tasks were employed to measure subjects' reading comprehension of texts in the target language: written recall protocols (written in L1) and multiple-choice questions (written in FL). The subjects' responses to these two different tasks were analyzed both qualitatively and quantitatively. The main objective of this study was then to address the following research questions:

1. Do subjects' familiarity versus unfamiliarity with the contents of the reading passage determine
 - a. their mean scores on the recall task?
 - b. the pattern of elaborations and distortions expected to appear in their written recall protocols?
 - c. their mean scores on the multiple choice task?
 - d. their performance on inference, generalization, and factual type of questions?
2. Do multiple choice tasks differ from written recall protocols in the nature of information they provide regarding the readers' comprehension of a text?

2. METHODS

2.1. Subjects

The subjects who participated in the study were students enrolled in a one-year intensive English program in the department of Foreign Languages at Gaziantep University, Gaziantep, Turkey. They were the prospective students of the Faculty of Engineering and the department of English Language and Literature. Since the medium of education was in English in the departmental studies of these subjects, it could be assumed that they were instrumentally motivated to learn English for academic purposes. At the beginning of the academic year a total of 360 students were randomly placed into 16 mixed-ability groups, with 15 groups comprising science-oriented engineering students, and two groups of language-oriented literature students. Throughout the academic year a total of 48 students dropped out of the program. The remaining number of students for whom complete data were available was 264.

The intensive language program consisted of six 50-minute lessons daily. By the time the study was conducted, subjects had been taught English for six and a half months in the department.

2.2. Instruments

Texts: Subjects in this study received two reading passages describing the marriage customs of two different cultures, one of them relating to the subjects' own culture (i.e. familiar text) and the other to a less familiar culture (i.e.

unfamiliar text). The Turkish culture was depicted in P.J. Magnarella's "Tradition and Change in a Turkish Town" [14] while the Ulithian culture was described in "Cultures Around the World: Five Cases" by G.L. Spindler [15]. To facilitate between-text comparison, the texts were matched and modified in their semantic, conceptual, and syntactic complexity. The length of the passages were determined both by the density of ideas in the texts and the time restrictions on the reading sessions.

The teachers of all the groups were asked to read the texts and underline the vocabulary items which had not been introduced to their students in the textbooks. This information determined 37 unknown words in the Ulithian text, and 24 in the Turkish text. The latter was kept unchanged while for the former text 12 words were substituted by their more familiar synonyms, and confirmed by the same teachers. The two texts were also compared from a syntactical point of view. Flesch's [16] readability formula applied to both texts yielded a grade level of 10.77 for the Ulithian text and 10.73 for the Turkish text. Both texts covered the same amount of information in almost the same number of paragraphs. The revised version of texts are introduced in Appendix A.

Tasks: Following Lee [10], Shohamy [11] and Gordon and Hanauer [12], the current study made use of two assessment tasks - multiple-choice questions and written recall protocol - to investigate whether the nature of information obtained from one task would be different from or similar to that obtained from the other.

Multiple-choice items for each text consisted of eight questions, five of which were inference questions, two were factual questions, and one being a generalization question. Factual questions required subjects to recognize verbatim factual details; the inference items entailed the inferring and interpretation of underlying meanings and relationships; and the generalization items required subjects to evaluate the whole passage or large portions of the text and to reach a conclusion regarding its contents [17].

The recall procedure is one of the most commonly used reading comprehension assessments because it is believed to be "the most straightforward assessment of the result of the text-reader interaction" [8]. Second language research using this technique differ in design and findings. Lee [18] found that passage recall was significantly better when done in the subjects' native language than in the target language, and that pre-reading instructions could affect recall. Free recall places a heavy demand on the subjects' written production skills, and reproducing the written recall in the target language can distort the research findings. Therefore, subjects in the present study were asked to reproduce their recalls in their native language. Pre-reading instructions were also provided in their L1.

2.3. Data Collection Procedures

Before data collection, the researcher met with the 17 teachers who would be administering the tests to the groups to instruct them on the methodology and timing. The reading activity and the following assessment tasks were administered to the subjects in two consequent 50-minute lessons. The order of the texts to be read was counterbalanced, so that while half the students received the familiar text first the second half worked on the unfamiliar text and vice-versa. The subjects were instructed to read the passage in 20 minutes and to answer the following multiple choice questions by looking back to the text whenever necessary. As soon as these questions were answered, teachers collected the passages together with the question sheets. Next, subjects were given out a blank page each with the titles of the reading passages typed on each side together with instructions to write down in Turkish everything that they understood and remembered from the passage. Since the texts were too long for total verbatim recall, it was expected that the subjects would recall and write down what they considered to be the most important aspects of the passages.

2.4. Data Analysis

During the study, subjects read passages in English but were asked to write their recalls in their native Turkish. Therefore, for a quantitative analysis of their recalls it was necessary to parse the Turkish translation of the original texts into idea units. One problem with studies of this kind has been to decide how far to go in analyzing the textual propositions. Meyer [19] argues that diagramming a passage to its fullest extent is not essential for most studies in education and psychology. In this study, it was sufficient for the author to know whether a subject remembered the basic notions in the text. Analyzing the arguments of propositions in a lengthy passage of 708 words would increase dramatically the number of units to score and make the scoring almost impossible with several hundred students. Even when working with larger units of analysis, one recall took approximately 40 minutes to score.

Therefore, the scoring procedure was modified in a way that an idea unit which carried only one bit of information significant for the analysis corresponded to certain grammatical units. The following are examples to the verbal, noun, clause, and phrase structures which were considered as single idea units.

Verbal structures: V (gelir - comes); V + to + V (ayak uydurmaya çalışmak - try to adapt);

V + Adj. (mutlu sayılırlar - are considered happy);

V + Ving (balık tutmaya giderler - go fishing)

Noun structures: Nouns, compound nouns, noun phrases.

Clause structures: adverbial clauses (starting with as, while, as soon as, after, etc.); noun clauses (starting with "It is emphasized that..., It is believed that...", etc.);

Phrasal structures: prepositional phrases, adverbial phrases, adjective phrases, participial

phrases, gerund phrases, and infinitive phrases.

In compound sentences expressing cause-effect, condition-result, and purpose-action relationships, each clause was treated as a complete sentence in itself and parsed as such.

The translation of the familiar text parsed in this way yielded 186 idea units while the unfamiliar Ulithian text amounted to 214 idea units. Each subject's recall score consisted of his or her correct recapitulation of the propositions in the texts. Although the possible range of propositions was 0 - 186 for the familiar text and 0 - 214 for the unfamiliar text, it was not expected that readers would score anywhere near that because of the time limit on the writing task. Each recall was scored by two independent raters and the inter-rater reliability was .95.

Several research studies give support to the argument that texts are never totally explicit and must be completed by the reader. Steffensen and Joag-Dev [20], for example, assert that when reading texts relating to their own culture, readers produce elaborations to supplement the explicit textual information. They also refer to distortions which could either involve an alteration of the explicit text, such as reversing the order of events or the order of words, or an addition to the text information which is judged to be inappropriate. In a qualitative analysis of written recalls, the investigator expected to find elaborations, which could be idea units that either did not exist in the passage or were an expansion of an existing element; and distortions, which were recalled units that either contradicted an existing piece of information - culturally inappropriate (CI) distortions - or distorted an existing piece of information in the text - culturally appropriate (CA) distortions.

3. RESULTS AND DISCUSSION

3.1. Written Recall Protocols

Table 1 presents mean performance in the percentage of recalled idea units and the number

Table 1 - Mean Performance as a function of Text Type (n=255)

Text type	Mean % of Idea Units Recalled	Elaboration (SD)	CI type distortion (SD)	CAtype distortion (SD)
Text 1 (fam.)	20.12 (10.30)	12.16 (9.75)	1.14 (2.02)	0.77 (1.35)
Text 2 (unfam.)	10.91 (7.36)	5.81 (5.67)	3.23 (4.07)	2.07 (2.32)
Results of ANOVA	F = 127.82*	F = 80.35*	F = 53.65*	F = 59.79*

*p<.0001

of elaborated and distorted idea units for both passages. The first research question related to whether performance on the recall task would be determined by the amount of information in the text which conformed to the readers' existing knowledge of the topic (i.e. their familiarity.) The results of the study showed that following their reading of the familiar text, subjects correctly recalled an average of 20.12 percent of the total idea units, which was twice the percentage of recall for the Ulithian text (i.e.10.91 percent). The difference between these means was significant at p<.0001.

The second research question was whether subjects would differ in the number of elaborations and distortions they make across texts. The values in the "elaboration" column of Table 1 proved text type to be a significant factor affecting percentage of elaborations in students' written recalls Compared to the unfamiliar text, students made more than twice as many elaborations in their recalls of the familiar text (i.e.5.81 vs. 12.16%, respectively). This could suggest that students were able to infer relationships between ideas and events, reach conclusions, and elaborate on the possible causes and results of events in the familiar text.

Closely related to elaborations are cases of distortions observed in subjects' recalls. Distortions were expected to appear in the recalls of the foreign text due to students' unfamiliarity with the foreign culture. As can be seen in Table 1, the number of distortions made in recalling the familiar text was on the average 1.91. idea units, whereas for the unfamiliar text it was 5.30 idea units. Such a significant difference in the amount of erroneous recalls may point to a me-

aningful relationship between distortions and the subjects' foreignness to the contents of the passage. Qualitative analysis of subjects' written recall protocols in a later section will provide the reader with more specific information regarding the nature of these elaborations and distortions.

3.2. Multiple Choice Questions

The average success of subjects on the eight multiple choice questions following the Ulithian passage was 31 % whereas the correct answers for the Turkish passage amounted to 54%. The questions constructed for this task, which consisted of factual, inference, and generalization type of questions, were presented to several colleagues and an American academician expert on reading [17] for confirmation for questions types prior to application. It was revealing to examine the distribution of the correct answers to these three categories, as illustrated in the following table.

Table 2 - Distribution of Correct Answers to Multiple Choice Items

TEXT1(familiartext)		TEXT 2 (unfamiliar text)	
<u>Factual Questions</u>			
#2 -	58%	#4 -	61%
#4 -	67%	#5 -	44%
Ave.	62.5%	Ave.	52.5%
<u>Inference Questions</u>			
#1 -	68%	#1 -	32%
#3 -	61%	#2 -	18%
#5 -	21%	#3 -	22%
#6 -	39%	#6 -	33%
#7 -	54%	#7 -	38%
Ave.	48.6%	Ave.	28.6%
<u>Generalization Questions</u>			
# 8 -	62%	# 8 -	31%

The two factual questions for the Turkish text were answered correctly by 62.5 % of the subjects while the same type of questions for the Ulithian text were answered by 52.5% of the subjects on the average. This means that when texts could be referred back to, text familiarity was not a determining factor in the subjects' ability to locate explicitly stated information.

With inference questions, success seemed to vary according to the degree of inferring required by the question itself (such as 21 % correctness for question 5 versus 68 % correctness for question 1 in the Turkish text).

However, on the whole, subjects' success with inference questions following the Turkish text seemed to surpass their performance on the Ulithian text (48.6 % versus 28.6%, respectively). Inferring underlying meanings and relationships in a text seems to be a skill which can be demonstrated more successfully when reading texts with a familiar content.

With generalization questions, although only one item for each text, subjects showed a similar tendency for higher success with texts of familiar content (62%) and lower success with texts of unfamiliar content (31%).

3.3. Qualitative Analysis of Subjects' Responses

Based on her survey of L2 reading research, Wolf [2] argues for the shortcomings of the multiple choice questions, in that many of the test items can be answered correctly without reading the passage and that such tests do not require careful reading of the passages. She also observed that subjects match words and phrases from the passage with words and phrases in the item stem and options. This argument seemed to be supported by the findings of the present study. Question 2 for the Ulithian text, for example, writes

Ulithians are not allowed to marry

a. if they have not reached a certain age.

b. if they carry on premarital sex relations.

c. if they are selfish and lazy,

d. if they do not find a handsome mate.

Almost half the students (43 %) marked "c". It is highly suspected that these subjects automatically matched the words "selfish and lazy" with the sentence in the text saying "If a man were to avoid marriage, it would be said that he is both selfish and lazy". Although the syntactic structure of the sentence was not beyond their linguistic capacity, subjects did not check the relevance of the item to the meaning in the text.

However, the results of this study would reject the assumption that multiple choice tasks do not require careful reading (at least not in all cases). Question 5 for the Turkish text (relating to paragraph 6) required the reader to infer the fact that only womenfolk take part in the engagement ceremony because there is no mention of the male fiancée's presence (only 21% realized this), but the careless readers who marked "d" (48%), which said that at the engagement ceremony "the women talk about the bridewealth and the trousseau", overlooked the information that the women had already discussed and settled matters related to bridewealth and trousseau prior to the engagement ceremony. All inference questions, by nature, require a careful reading on part of the student.

Question 8 for the Turkish text referred to the importance of the marriage institute in that culture and required subjects to reach a conclusion about the possible reason of this. Although the majority marked the correct option (62%), 22 % of subjects marked option "d" which said marriage was important for the Turkish community because "it is difficult for find honorable families to be one's future in-laws". These subjects may have used their world knowledge rather than linguistic knowledge to reach the conclusion that since a considerable amount of time and effort is spent in searching for "honorable and reputable families" this might be something not too easy to find. However, this answer did not

capture the cause-effect relationship underlying the stem and its option. It might be of relevance here to note that none of the recalls included information on the difficulty of finding honorable in-laws, but quite a number mentioned the importance of finding such people. When the student is allowed to relate his understanding of the passage in the absence of any written probes he tends to avoid reproducing such counter or illogical arguments.

Gordon and Hanauer [12] carried out a study to investigate the interaction between meaning construction and testing tasks. Based on the findings of this study, they concluded that testing task functions as an additional information source which influences the development of the mental model of the text. Multiple choice items, in particular, provide more information for the test taker to consider and therefore have a great potential for influencing the ongoing procedure. Such an assumption would automatically lead one to expect to see propositions which were embodied in the multiple choice questions, especially those in the inference and generalization items, to appear in the subjects' recall protocols. Although this study was not carried out with the purpose of seeing how much of the information induced by one task was transferred to the contents of another task, it was yet interesting to note that some ideas did seem to transfer from the earlier task while some did not appear at all. For the Ulithian text, for example, a number of propositions from question 7 appeared in students' recalls, as "boys are an economic advantage to the family", "they [Ulithians] want their children to be strong", or "they want to have their first baby a boy". For the Turkish text, information from question 7/d was widely expressed in the recalls; namely that "the marriage ceremony takes place on Sundays because many people are free to join the ceremony". One explanation for this could be that textual information which conforms to the students' existing world knowledge, whether in agreement with the reading passage or not, remains with the student and transfers from one testing instrument to the other.

However, examples could also be given for propositions that showed no transfer at all. For the first question following the Turkish text 18 percent of the readers marked the option that said "girls usually have difficulty in deciding whom to marry", whereas this proposition did not appear in any of the recall protocols. Similarly, the first question regarding the Ulithian text asks whom a Ulithian would prefer to marry, and while 23 percent of the subjects marked "someone from the chief's family" this piece of information did not appear at all in the subjects' recalls. Within the same frame of logic, it could be said that a task-induced proposition did not transfer across different measurement tasks unless it coincided with the subjects' existing system of world knowledge. The written recall task, by nature, does not influence readers' comprehension of a text because it does not provide any additional clues to the meaning of the passage [2].

Subjects' success with the inferencing and generalization questions in the multiple choice task may seem to be moderate (see Table 2) but this should not be taken to mean they lacked this skill. Most recall protocols contained at least one of inferencing, generalization or summarizing propositions that did not exist in the text. Examples are:

(R: Recall translated into English; T: Text)

R: In short, Ulithians are a community where people have no sex problems, and just eat, drink, and sleep.

R: To conclude, members of the opposite sex go through a flirting stage and later decide to marry.

R: Two types of marriages are mentioned in the passage: the first is the marriage to more than one women, and the second is the premarital sex relations and the marriage that follows this.

R: The passage describes the Turkish custom of arranged marriages.

The recall protocols also allowed the researcher to observe subjects' interaction with the text on culture-related matters, as exemplified below:

R: If tea is served too soon, it implies that the daughter of the house is too young for marriage.

T: (If tea is served too soon, the guests feel their visit is being rushed)

R: Because the family elders wish to choose a bride who will conform to the new family, her selection was not left to the groom only.

T: Because a new bride joined the household of her husband's parents, her selection was important to all its members, not just to the groom.

Personal comments and interpretations also appear, as in the following example:

R: Such [Ulithian] marriages would not last long.

R: The girl's dowry belongs to the boy's family.

R: They are ahead of our times in their traditions.

Elaborations on the Ulithian text were usually based on cross-cultural transfers from the readers' own cultural experiences. These elaborations did not relate at all to any information in the text and sometimes even contradicted it, but they could be traced back to the subjects' own culture.

R: Man is the head of the family

R: After marriage, the woman has to support her husband and be obedient.

R: If families cannot get along, they separate and the food gifts are sent back.

Culturally inappropriate (CI) distortions relating to the Ulithian text also gave signs of cultural intrusions from the subjects' culture:

R: The rich cannot marry the poor.

R: The parents of the couple to marry want to make all the decisions

With any set of multiple choice questions, subjects' understanding of a text is measured through a limited number of tester-determined questions, in which case individual differences in understandings and misunderstandings, varying interpretations and inferences are all concealed for the sake of uniformity and easiness of scoring. The recall protocols, however, enable researchers to determine some sources of difficulties through subjects' reconstruction of what they understand. This will be discussed under the following heading.

3.4. Some Sources of Difficulties

The sources of difficulties were investigated through an analysis of the incorrect recalls, which were called distortions. When a proposition did not appear in any of the recalls, the source of the problem was suspected to be on a sentential, a clausal, or a lexical level. There were sentences in both texts which were altogether omitted from all the protocols, such as the following examples:

T: These two important considerations are accompanied by a third one involving the interfamilial alliance or acquisition of relatives that marriage entails.

(problem words: considerations, accompanied by, interfamilial alliance, acquisition of relatives, entails)

T: The preliminary steps protect the men from the loss of honor which a face-to-face refusal would bring. (preliminary steps, loss of honor, face-to-face refusal)

Sometimes it was only sentence fragments that were omitted, such as the underlined parts of the following examples:

T: This package is only a part of a complex exchange of gifts between members of the two kin groups.

T: Because honor and shame, it is vital to choose in-laws who will not dishonour one's offspring as well as one's household and lineage.

T: One might take the skeptical view that the desire for children is motivated by economic considerations, but this would ignore the obvious delight that a man and a woman take in having young ones about them.

With omissions it was hard to make a definite statement of whether the sentences or sentence segments were excluded due to vocabulary difficulties, or syntactic difficulties, or both. It is also possible that subjects believed these parts to be unimportant for the message of the whole. However, more convincing evidence to miscomprehension due to syntactic difficulties could be found in what subjects included in their recalls. Presented below are some examples of recalls which showed that subjects had syntactic difficulties:

R: It is important to raise children both economically and physically strong.

T:the desire for children is motivated by economic considerations, but

R: According to Ulithian traditions, a young girl and boy cannot carry on premarital sex relations until they have acquired the necessary qualities for a marriage.

T: Marriage is permitted upon reaching adolescence but is usually postponed for a few years during which the young man and woman is free to carry on premarital sex relations.

R: When searching for a girl, the daughter of a patriarchal family was looked for because she would become a member of their family.(intersentential error)

T: The traditional method of selecting spouses was functionally consistent with the ideal patrilocal extended household. Because a new bride joined the household of her husband's parents, her selection was important to.....

The above examples showed that subjects had difficulties on the sentential and intersentential level. Therefore, they tended to break the longer sentences into their smaller segments, which in turn resulted in incomplete and wrong meanings.

Both texts were written as descriptive listings of events leading to marriage. However, the discourse structure of the Turkish text led to two major problems that could be detected in the subjects' recalls. Firstly, information in the first two paragraphs culminated in the third, where the paragraph began with "These two important considerations are accompanied by a third one". The first two considerations were not signaled by any discourse markers, and all three considerations were later to be tied to the conclusion "Therefore, marital decisions could not be left to young individuals." However, readers did not realize the cohesive structure among the first three paragraphs which was achieved by the cause-effect relationship. Therefore, rather than listing the three major reasons that lead to the intensive family involvement in the marriage decisions, they referred separately to one or two reasons of family involvement without expressing the causal ties between them. This conclusion was confirmed by the omission of the first sentence of the third paragraph and the cohesive ties that signaled a connection between the ideas. As Cohen suggested [21], EFL learners tend to focus on isolated facts and details, whereas an in-depth reading would enable them to understand textual relationships, such as cause and effect. They remain unable to synthesize information from different levels of a passage, which would require integrating details as well as main ideas. The written recalls of the native text mostly centered on the ceremonial parts of the procedure like "seeing", the sending of intermediaries, the engagement ceremony, and the wedding.

Another type of problem occurred in relation to the different stages of the marriage procedure in the Turkish text. There are several gat-

herings of women which start with the informal visits described in the fourth paragraph, followed by the first formal meeting in the fifth paragraph, which in turn lead to the coming of the male and female intermediaries to ask for the father's consent. The sixth paragraph gives a detailed account of the next gathering as the engagement ceremony. Subjects were confused with the number of these visits, so that in some of the protocols the contents of the engagement ceremony were repeated for the "seeing" event; or, the "seeing" ceremony was confused with the initial informal negotiations of the women; or, "the financial particulars" were discussed in the engagement ceremony; or, the ring was placed on the girl's finger as part of a pre-engagement, and so on. These are all possible variations of the same cultural events happening in different parts of the country. Subjects' existing background knowledge helped to both facilitate and complicate their reconstruction of the text. The Ulithian text did not yield such problems. For this text the frequency of recalls were higher for those propositions which appeared in the initial position of paragraphs as the introductory statement to the contents of the paragraph, and there was a more even distribution of propositions recalled from different parts of the text.

4. CONCLUSION

The results of the quantitative and qualitative analyses of data lead to the conclusion that a combination of factors need to be taken into account when discussing a subject's performance in an L2 reading comprehension task. First of all, the reading passage needs to be examined carefully from the point of view of its potential linguistic difficulties, subject's prior knowledge related to its content, and its discourse structure. Second language readers in this study had difficulty coping with parts of the text that contained difficult vocabulary and/or syntax regardless of whether they were familiar with the topic or not [22]. Parts of the text that contained such difficulties were either omitted from recall protocols,

or were recalled incorrectly. It was usually the longer sentences that posed these difficulties. A second major finding of this study was that subjects failed to take notice of intertextual relationships within and between paragraphs and therefore did not demonstrate a global understanding of the text.

Two assessment tasks were used in this study: multiple choice questions in the target language and recall protocols in the native language. While the multiple choice task measures subject's ability to comprehend the information in the question stem and its options in addition to understanding the textual information, the recall protocols emphasize the importance of comprehending passage information in that subjects must construct what they understand. The data obtained from these two tasks gave evidence to the earlier stated hypotheses that readers who are familiar with the content of the passage understand and accurately recall more than do readers less familiar with text content [20]; perform better on both objective tests and written recall [23]; and make more correct inferences after reading texts in their own culture [24]. In their interaction with the text subjects made more elaborations in response to their reading of the native culture and more distortions based on their reading of the foreign culture. Subjects on the whole had more difficulty in answering inference type questions than the others, irrespective of the text familiarity factor, but this should not be taken to mean that lacked inferencing skills. Most of the elaborations (and even distortions for that matter) found in their recalls consisted of propositions inferred from or generalized on the contents of their reading. Some of the difficulty students had in answering multiple-choice questions could also be attributed to the fact that they were written in the target language.

There was a very little amount of transfer of information from the contents of the multiple choice task into the subjects' recalls. If testing task had functioned as an additional source of information during subjects' mental construction

of the text [12], this would have been possible. It could be argued that task-induced inferences and ideas tend to be suppressed in the subjects' memories in favor of self-generated inferences.

A careful analysis of the recall protocols in this study revealed that there were linguistic and conceptual difficulties experienced by all foreign language readers while some were specific to the individual. However, merely focusing on the quantitative analysis of the recall protocols does not provide one with sufficient information about the quality of recalls. With such practice information from extratextual sources, such as those that lead to elaborations, distortions, and cultural intrusions would be totally ignored. It is advisable, therefore, to use the written recall procedure as a testing device in diagnostic tests in order to obtain detailed information about students' strengths and weaknesses and the specific problems they encounter when reading. The use of a combination of tasks is also advisable because each task yields information which complement the information from the other.

5. LIMITATIONS

There are several limitations to this study. First of all, the researcher did not pre-determine the nature of the subjects' existing background knowledge related to the contents of the two passages. The second limitation relates to the use of the immediate free recall protocols which are believed by some to be examining only the product of comprehension and not the process of comprehending. Therefore, subjects' protocols may not have reflected all the mental activities that have taken place in their head, and may be limited in their value as indicators of the comprehending process. And thirdly, the type of questions used for the multiple choice task, such as factual, inference, and generalization, do not exhaust all the possibilities that could be included into an assessment of this kind. Moreover, the limited number of these questions still leaves a

lot of information related to the text untapped. However, within the time limits and aspirations of the study eight questions for each text was all that could be included into the multiple choice task.

BIBLIOGRAPHY

- [1] Bernhardt, E.B. "Reading development in a second language: Theoretical, empirical, and classroom perspectives." Ablex Publishing Co., (1991)
- [2] Wolf, D.F. "Issues in reading comprehension assessment: Implications for the development of research instruments and classroom tests". *Foreign Language Annals*, 26: 322-331. (1993).
- [3] Farr, R., Carey, P., and Tone, B. "Recent theory and research into the reading process: Implications for reading assessment". In J. Orasanu (Ed.), *Reading Comprehension: From research to practice*, Hillsdale, NJ: Lawrence Erlbaum Associates, (1986).
- [4] Carrell, P.L. "The effect of rhetorical organization on ESL readers". *TESOL Quarterly*, 18; 441-469, (1984).
- [5] Carrell, P.L. Awareness of text structure: Effects on recall. *Language Learning*, 42: 1-19, (1992).
- [6] Adams, M.J., and Collins, A. "A schema-theoretic view of reading". In H. Singer, and R.B. Ruddell (Eds.), *Theoretical models and processes of reading, 3rd ed.*, Newark, Delaware: International Reading Association, (1985).
- [7] Carrell, P.L. "Introduction: Interactive approaches to second language reading". In P.L. Carrell, J. Devine, and D.E. Eskey (Eds.), *Interactive Approaches to Second Language Reading*. Cambridge: Cambridge University Press, (1988). *comprehension in intermediate German*".
- [8] Johnston, P.H. "Reading comprehension Assessment". Newark, Delaware: International Reading Association, (1983).
- [9] Chen, Q. and Donin, J. "Discourse processing of first and second language biology texts: Effects of language proficiency and domain-specific knowledge". *Modern Language Journal*, 81: 209-227, (1997).
- [10] Lee, J.F. "Comprehending the Spanish subjunctive:

- An information processing perspective". **Modern Language Journal**, 71: 50-57, (1987).
- [11] Shohamy, E. "Does the testing method make a difference? The case of reading comprehension". **Language Testing**, 1: 147-170. (1984).
- [12] Gordon, C.M. and Hanauer, D. "The interaction between task and meaning construction in EFL reading comprehension tests". **TESOL Quarterly**, 29: 299-323, (1995).
- [13] Bernhardt, E.B. "Three approaches to reading **The Modern Language Journal**, 62; 111-117, (1983).
- [14] Magnarella, P.J. "**Tradition and change in Turkish town**". New York: Halsted Press. (1974).
- [15] Spindler, G.L. "**Cultures around the world: Five cases**". New York: Holt, Rinehart,
- [16] Flesch, R. "A new readability yardstick". **Journal of Applied Psychology**, 32: 221. (1948).
- [17] Perkins, K., Brutton, S.R., and Pohlman, J.T. "The effect of second language proficiency on second language reading comprehension as measured by questions from different cognitive levels". **Per Linguam**, 7: 49-57. (1991).
- [18] Lee, J.F. "Background knowledge and L2 reading". **Modern Language Journal**, 70: 350-354 ,(1986).
- [19] Meyer, B.J.F. "Identification of the structure of prose and its implications for the study of reading and memory". **Journal of Reading Behavior**, 7: 7-47, (1975).
- [20] Steffensen, M.S., and Joag-Dev, C. "Cultural knowledge and reading". In J.C. Alderson, and A.H. Urquhart (Eds.), **Reading in a Foreign Language**. New York: Longman. (1984).
- [21] Cohen, A.D. "On taking language tests: What the students report". **Language Testing**, 1: 70-81, (1984).
- [22] Kim, S. "Types and sources of problems in L2 reading: A qualitative analysis of the recall protocols by Korean High School EFL students". **Foreign Language Annals**, 28: 49-67, (1995).
- [23] Floyd, P., and Carrell, P.L. "Effects on ESL reading of teaching cultural content schemata". **Language Learning**, 37: 89-108, (1987).
- [24] Steffensen, M.S., Joag-Dev, C., and Anderson, R.C. " A cross-cultural perspective on reading comprehension". **Reading Research Quarterly**, 15: 10-29. (1979).