ЛИНГВИСТИКА

UDC 81'243+ 81'33
DOI: 10.17223/19986645/73/1

Esmat Babaii[1], Tatiana M. Permyakova[2], Ekaterina V. Pozdeeva[2]

[1] Kharazmi University (Tehran, Iran)
[2] Higher School of Economics (Perm, Russian Federation)
E-mail: babai@khu.ac.ir / tpermyakova@hse.ru / epozdeeva@hse.ru

Strategies in Performing a Multi-Level C-Test: Applying Think-Aloud Protocols

The aims of this study were: 1) to identify the cognitive strategies that Russian EFL learners apply while taking a multi-level C-test; 2) to examine the correlation between the level of complexity of a C-test and the frequency of the usage of the cognitive strategies. The Think-Aloud Protocols method allowed to reveal seven cognitive strategies (predicting, summarizing, questioning, making connections, re-reading/using fix-ups, identifying a problem, and reflecting). The results of chi-square tests indicate that texts with different complexity levels can trigger different type of processing used by the test-takers.

Keywords: *c-Test; cognitive strategies; language proficiency assessment; think-aloud protocols*

Introduction

C-tests are gap-filling tests that consist of several short texts in which parts of words are missing. They are widely used to assess general language proficiency for purposes of placement, screening, or provision of feedback to language learners. Various studies have revealed that C-tests are integrative and highly reliable and valid measures of general language proficiency [1–3]. This study is motivated by the work of Babaii and Moghaddam [4], with the underlying assumption that the C-test at different proficiency levels triggers the use of different test-taking strategies on the part of learners. We seek to investigate how the shift in the text difficulty level of the C-test input (proficient, intermediate, and beginner) influences the strategies employed by test-takers. We attempted to answer this question through applying an introspective think-aloud protocol (TAP) method.

Literature Review

The C-test was designed as an instance of reduced redundancy testing and was claimed to be more representative in terms of assessing language proficiency than multiple cloze-tests or cloze-elide tests [3]. A typical C-test consists of four or five short texts where the second half of every second word is deleted

(the rule of two). The gaps have to be filled in by second/foreign language test-takers. Each passage has approximately 20 gaps. The first introductory sentence is left intact while the rule of two applies to the rest of the passage.

The C-test is believed and confirmed to be a well-established measure for L2 proficiency level, in different language contexts, for example, Estonia [5, 6], Germany [7], Hungary [2], Thailand [8], Iran [9], and China [10]. To the best of our knowledge, the only published research recruiting a Russian-speaking sample is the one conducted by Drackert and Timukova [11] where the C-test was found to be biased towards heritage language learners and therefore was not deemed satisfactorily valid.

The C-test is a highly reliable and valid instrument [7], as it also correlates well with other proficiency tests such as MTELP – Michigan Test of English Language Proficiency [12], MCT – a multiple-choice achievement test, TOEFL – Test of English as a Foreign Language [9], or a Cloze test [3]. The most notable contradictory result, perhaps, is reported by Jafarpur [13] who investigated the varieties of the C-test by applying different deletion procedures. The results of the study revealed that different versions of the C-tests did not bear similar characteristics, and findings refuted the claims on C-testing as an adequate measure of general language proficiency. Hastings [14], in response, provided a detailed analysis of Jafarpur's methodology vis-a-vis the principal rules of the C-test and concluded that his controversial findings were due to flouting the rules.

Furthermore, when it comes to the support of applicability of the C-test to various specific contexts, the results are mixed. Dörnyei and Katona [2] believe that the C-test is applicable to different proficiency levels. Muller and Daller [15] used C-test with correlation to IELTS ($r = 0.509$, $p < .001$) to measure whether these tests are appropriate for the medical professional context. Similarly, Prediger et al. [16] found that language proficiency is the background factor with the strongest connection to mathematics achievement among all social and linguistic background factors. The study by Daller and colleagues [17] corroborates the validity of the C-test in examining the fluency and overall oral proficiency among bilingual speakers, with correlations ranging from .31 to .63 for different determiners of proficiency. On the other hand, Mashad's study [12] claims that the C-test does not prove to behave consistently with examinees of different proficiency levels; in particular, it cannot differentiate participants of lower and upper intermediate levels. To solve the problem, modifications to C-tests were made to increase and decrease the difficulty of the test by manipulating the size and the distribution of gaps in order to be able to generate C-tests with the desired difficulty levels [18]. These findings were consistent with the reliability of test scores in a modified C-test procedure [19].

Despite the abundance of research on the C-test and its psychometric characteristics, there is a scarcity of research regarding possible interactions between text characteristics and the nature of processing in the C-test completion. One of the works that attempts to investigate whether difficult test tasks can trigger more macro-level processing in C-testing is by Babaii and Moghaddam [4]. Their research involved manipulating degrees of syntactic complexity and ab-

straction as text variables and presenting the chosen texts in two formats: with or without clues regarding the number of missing letters. Out of the sample of 119 students, 36 participated in the retrospective think-aloud phase of the study. The think-aloud protocol analysis revealed that employing texts with more syntactic complexity and abstraction, along with eliminating clues with respect to the number of missing letters can result in more difficult test tasks which seemed to encourage a more frequent use of macro-level processing on the part of test takers.

Several studies focus on the interplay between text difficulty, learners' language proficiency level and the type and frequency of (meta)cognitive strategies learners use for processing and understanding. Kaivanpanah and Moghaddam [20] investigated how different levels of L2 reading proficiency may affect the readers' inferencing behavior with respect to the level of success they achieve as well as the types, combinations, and frequencies of knowledge sources they use when inferring meanings of unknown words. They found that differences in L2 reading proficiency level had a profound impact on inferencing success. However, the contribution of the knowledge sources used by the participants was found to be partial and restricted. Their findings partially confirmed the results of previous studies [21] on the contribution of topic familiarity and text difficulty on the test takers' performance in reading comprehension tests in language for specific purposes.

Similarly, Zhang [22] contends that EFL proficiency is linked with language learners' metacognitive knowledge or awareness of strategies. Using TAP, Zhang [23] employed a 28-item survey of reading strategies to explore relationship between metacognitive awareness and reading-strategy use of Chinese high school students. The strategies were classified into three categories: global, problem-solving, and support. The high-proficiency group outperformed the intermediate group and the low-proficiency group in two categories of reading strategies: global and problem-solving; but no statistically significant difference was found among the three proficiency groups in using support strategies.

A number of introspective studies examined the effect of EFL learners' reading proficiency on the nature of lexical inferencing they employed while reading in English. Learners were found to use the knowledge sources and contextual cues differently; and the level of success and their rate of learning and retention of the inferred target words varied [24]. Reading proficiency in L1 is a contributing factor to general proficiency measured by C-testing and other text-dependent tests [25]. Yet, it seems that learners need to establish some knowledge of an L2 per se before they can successfully draw on their L1 reading ability to help with L2 reading in certain tasks like C-test completion.

An intriguing contribution to this line of research was done by Soemer and Schiefele [26] dealing with the question of how the difficulty of a text could affect a reader's thinking including but not limited to task-unrelated thinking (mind wandering) during reading. In the study, participants read three texts rated as easy, moderate, or difficult in terms of readability and cohesion. During reading, participants were required to indicate whether they were engaging in mind wandering. The results showed that reading difficult texts increased mind wan-

dering and this increase partially explained the negative relation between text difficulty and comprehension.

Among introspective techniques used to detect L2 learners' mental processes, Think-Aloud Protocol (TAP) method is a feasible procedure alleviating the need for utilizing highly sophisticated equipment to explore the cognitive processes which are not directly observable [27]. So far, there are few works on the accounts of test-takers' mental processes in the C-test [28]. Babaii and Ansary [9] as well as Babaii and Moghaddam [4] employed TAP to explore the nature of processing in the C-test where both micro- and macro-level processing emerged after the analysis of test-takers' TAPs. In another study, Babaii and Fatahi-Majd [29] investigated the test-takers' failed attempts to complete the blanks through retrospective verbal protocols. High and low performers were reported to experience diverse psycholinguistic obstacles and mechanisms. While the high performers struggled with retrieving the appropriate lexical items and overlooked delicate points of grammar, the low performers suffered most from breakdown in text processing and automatic restoration based on the local clues.

To summarize, the study of available literature on the C-test points to the paucity of research on the modifications of this test in terms of difficulty levels. Specifically, literature overview shows a gap in the introspective use of TAP with regard to strategies employed while C-testing. We assume that the C-test at different proficiency levels may trigger the use of different reading strategies on the part of test-takers. The research gains momentum considering the fact that Russian EFL learners remain an under-investigated sample in this area of inquiry.

Method

Participants and settings

A total number of 20 participants took part in this study. To ensure the sample's homogeneity, they were all selected from Higher School of Economics, aged from 20 to 22, with the average age being 21.3, both males (50%) and females (50%). The participants were Russian native speakers with the recognized upper-intermediate level in terms of the English language proficiency. To ensure they all have the same level of reading proficiency, the participants were pretested and had to undergo an EF Standard English Test (https://www.efset.org/english-certificate). This test provides the analysis of the test taker's reading skills, and its results are aligned with CEFR levels.

This study applied a qualitative method using the Think-Aloud Protocol. The participants were asked to complete the missing parts in the passages and verbalize as much as possible anything that went through their minds while they were doing the task.

Materials

To ensure the reliability and validity of the material, the passages presented to the participants were extracted from Reading and Use of English sections of the test papers provided by Cambridge Assessment English (Appendix 1). The C-test contained three passages, with the first one extracted from B1 Prelimi-

nary, the second one from B2 First, and the third one from C2 Proficiency tests. The texts' lexical complexity increased from B1 to C2, with the latter containing more long words (6 letters and more, 3 syllables and more) and fewer frequent words. The syntactic complexity also grew from B1 to C2 text, with the latter having more participial constructions and nominal phrases. Each passage contained 24 to 25 gaps and was chosen in such a way that the content did not correlate with the test takers' professional areas to eliminate the possibility of some participants having an advantage over others. The reliability index calculated through KR-21 for the C-test was 0.78. This seems a satisfactory index of consistency considering the homogeneity of test-takers in terms of language proficiency level which generally leads to under-estimation of reliability [30].

Procedure

Data collection started in September 2018 and lasted for 2 months. The C-tests were presented to each participant individually on paper task sheets, and the assigned time for completing three passages was 30 minutes. The researchers asked the participants to verbalize whatever came to their minds while they were filling in the gaps in each passage. The researchers recorded the participants' verbalized thoughts providing prompts when they forgot to "think aloud" by asking questions such as "Why do you use this particular word to fill in the missing part?" The recordings were later transcribed into protocols that were used to analyze the data and spell out the strategies which the participants chose to perform the task.

As the test takers verbalized their thoughts, they typically used sentences such as "I think this should be <…. because …>.", "I will re-read the first sentence", "I do not know exactly what it is going to be but I think …", etc. As Baumann and colleagues [31] stated, such sentences can serve as indicators of the following strategies: Predicting, Clarifying, Questioning, Making Connections, Re-Reading/Using Fix-Ups, Summarizing, and Commenting. Derived from reading comprehension techniques, strategies were identified by verbalized items, usually preceding the blank to be filled. The researchers analyzed the protocols to recognize what strategies the participants used to complete the C-tests.

Analysis of Data and Results

The participants' think-aloud protocols were studied to find out what cognitive strategies were used when the participants filled the gaps in the C-tests. The data were converted to several types of cognitive strategies: Predicting, Summarizing, Questioning, Making Connections, Re-Reading/Using Fix-Ups, Identifying a Problem, and Reflecting.

The Predicting strategy involved making use of clues while doing the test. For instance, while filling in the gap in "t_____ discuss rese_____ into", one of the participants stated, "to discuss… research. I think it's research here because discuss goes with a noun".

Another strategy which was used by the test takers was Summarizing. The participants applied it when they put some information from the passage into their own words. In case of "hoping t____ get ri_____", one of the participants

said, "the passage is about discovering gold in Alaska, so I guess people came there hoping to get rich. Yes, it's rich here".

The Questioning strategy means that the subjects asked themselves questions when they faced difficulty trying to find the answer. For instance, faced with "plen___ of dir____ daily exper_____", a participant said, "plenty of some experience… dir…. daily experience….What experience could they have?"

Making Connections was a frequently applied strategy when the participants tried to make a connection between what they were reading and what they have already read, or their background knowledge about the world. Faced with "no sin____ governing bo___", one of the test takers said, "I think I have heard the phrase governing body. It can be a governing body here. No single governing body".

The Re-Reading/Using Fix-Ups strategy, as the name itself suggests, involves re-reading the text and the intact co-text. Dealing with "such st_____ training", the test taker said, "st… training… I should re-read the whole text again to understand".

Another strategy, Identifying a Problem, was used when the participants had to deal with those gaps that they did not know how to fill. They usually verbalized this strategy saying "I am not sure", "I do not know what to fill it with" or any other phrase indicating they have faced a cognitive challenge. For instance, having to deal with "and enfor_____ of" one of the subjects said, "It is difficult. I am sure it is about laws but I can't think of any word".

Reflecting as a cognitive strategy was applied when the participants "turned back" to explain or interpret their gap-filling decisions. For instance, in case of "to outs____ humans", the test taker said, "Initially, I thought it could be outstand but then replaced it with outsmart because the text said that rats were clever and intelligent".

During the procedure, strategies were identified by verbalized items, which usually preceded the blank to be filled. Some of the filled blanks were not commented upon by the respondents. When the two strategies could have been interpreted as overlapping, for instance, Re-Reading and Reflecting, the researchers turned to verbal protocols for the supporting markers such as phrases "I need to re-read" or "I think … because …" to distinguish between the two.

The analysis revealed that the participants used seven strategies while performing the gap-filling task. However, the usages of the strategies differed: while some of them were applied intensively, the others were used less frequently. Moreover, the frequency of occurrence of these strategies varied in correspondence to the level of the difficulty of the passages (B1, B2, C2). Table 1 presents the data on how often strategies were applied in each passage; figures in **bold** are 5% above the average and figures in *cursive* are 5% below the average in each column.

When the participants had to deal with the passage of the B2 level, which was pre-tested as their baseline, Making Connections was the most frequently used strategy (26%). In case of the B1 passage, the participants applied it slightly less often (22%); however, when they did the C2 level test, only 10% of the gaps in the text were filled using this strategy.

T a b l e 1

Frequencies and percentages of strategies used across different levels of text complexity

Level	Strategies							
	Predicting	Summarizing	Making Connections	Questioning	Identifying a Problem	Reflecting	Re-Reading/ Using Fix-Ups	
B1	29 (13%)	59 **(27%)**	48 (22%)	*11 (5%)*	21 (9%)	*10 (5%)*	**41 (19%)**	N = 219
B2	**92 (21%)**	*33 (8%)*	**110 (26%)**	30 (7%)	45 (10%)	**89 (21%)**	*28 (7%)*	N = 427
C2	35 (10%)	46 (13%)	*34 (10%)*	**56 (15%)**	**81 (22%)**	56 (15%)	55 (15%)	N = 363
Total by strategy	156	138	192	97	147	155	124	1009

Predicting was close to Making Connections at the B2 level – the test takers applied it to fill in 21% of the gaps. In case of the B1 and C2 texts, the use of this strategy followed the same pattern as Making Connections did – the frequency of occurrence decreased to 13% and 10% respectively.

At the B2 level, Reflecting was used to fill in 21% of the gaps. However, when the test takers had to fill in the gaps in the B1 passage, they "reflected" on their performance much less intensively – only in 5% out of the total number of gaps. In case of the C2 text, the frequency of occurrence decreased to reach 15%.

The rest of the strategies were used less frequent at the B2 level. Re-Reading/Using Fix-Ups" was used to fill 7% of the gaps. When the participants took the B1 and C2 tests, they applied it more often – 19% and 15% respectively.

The sixth applied strategy at the B2 level was Identifying a Problem, which was used to fill in 10% of the gaps. At the B1 level, its frequency of occurrence did not change much – the participants stated they had to deal with a problem in 9% out of the total number of tasks. However, when they had to do the C2 text, they applied it to 22% of the gaps.

Questioning was the least frequently used strategy in case of the B2 and B1 tests (7% and 5%), while in case of the C2 test the frequency of occurrence grew to reach 15%.

The data indicated that Summarizing and Reflecting were the strategies showing the highest deviation from the average at the B1 level – sharply going up for Summarizing and sharply going down for Reflecting. At the B2 level the trend was reverse – a sharp decrease in using Summarizing and a significant increase in using Reflecting. At the C2 level, the strategies that yielded the greatest deviation from the average were Making Connections (10% in contrast to the average of 19.3%) and Identifying a Problem (22% as opposed to the average of 13.6%).

For example, reading a C2 passage, in the Identifying a Problem strategy, a test taker often commented on not knowing the word: "There is no // govern-

ment / there is no government of what? / Legislature / what is this word? // issues and informs, enforces issues and enforces laws to which any nations / what is this word? / there is no / there is no / why do I want to say 'there is no worry'? it does not fit here / I do not know / government of legislature issues and enforces laws to which any / what is the word? I do not know its meaning / the international legal order has no sinful / it is either sin or sinful or something else". While dealing with lower level passages, the frequency of the Identifying a Problem strategy decreases since correct responses are ensured by automatic processing of high frequency lexicon (e.g. "there is no worry").

These processes are parallel to the Making Connections strategy. For instance, a participant said: "issues and enforces laws to which // to which government / it must be here because I heard it before / issues and enforces laws // There is no // government of legislature issues and enforces laws to which any // to which any nations and subjects / it does not connect well // the international league order has no / has no // I want to say 'sinful governing', but it does not meaningfully connect // Governing and operates by agreement between states". Overall, the main reason for the shift in the use of strategies while reading C2 passages is inability to rely on automatic processing of familiar linguistic items and respondents' acknowledgment of this fact.

In order to check whether the differences in the use of strategies are statistically significant, three sets of chi-square tests were applied to the data. As it is reported in Table 2, all three comparisons are significant at $p < .000$ with a strong effect size calculated through Cramer's V [cf. 32].

T a b l e 2

Results of chi-square test for strategies by text levels

Index	Chi-square	df	p-value	Cramer's V (effect size)
B1 vs B2	89.84	6	.000	.373
B1 vs C2	73.02	6	.000	.354
B2 vs C2	97.93	6	.000	.352

The findings indicate that texts with different complexity levels can trigger different type of processing used by the test-takers.

Discussion and Conclusion

The aim of this study was to investigate the use of cognitive strategies of the test takers performing a task that required making decisions and solving problems. We assumed that the frequency of occurrence of the strategies involved in completing a C-test could vary depending on the text difficulty (in terms of the English language proficiency levels).

The results of the study revealed that each test taker applied a set of cognitive strategies, with some of them being used more intensively. The extent to which each strategy was used reflects the way each participant thinks and solves cognitive problems.

Another finding was that the frequency of the use of each strategy changed depending on the passage that the participant had to deal with. As each participant had the recognized upper-intermediate level (B2) of English, the first passage (B1) was quite likely to cause less trouble for them while the third passage (C2) was expected to be more linguistically challenging. We assumed that the differences in the language difficulty of the three passages would influence the extent to which the strategies would be applied. The results showed that the frequency of reverse occurrence of the strategies Summarizing and Reflecting distinguish between reading texts at B1 and B2 levels. C2 text reading is characterized by the increase of Questioning and Identifying a Problem strategies.

From the theoretical perspective, the findings might be useful in assisting to arrive at a more unified theory for test-taking strategies [33] since this study provided information about the way a group of Russian learners dealt with different difficulty levels in a text-dependent test. A meta-analysis on the contribution of L1 to the C-test completion could benefit from our study, which recruited a Russian sample in a TAP study of the C-test for the first time.

In line with Mashad's study [12], the results of this research add to the complications of tackling C-test difficulty and its interaction with test-taking strategies as the observed strategies do not clearly discriminate between different difficulty levels. One possible explanation for this is the different degree of involvement in extensive reading beyond the classroom for some of the participants [34]. It may also suggest that further control for the test-takers' characteristics such as individual differences, core demographics (males and females, age, educational background) and problem-solving traits is needed [35–38]. Nevertheless, at this point, the use and shift of strategies while performing a multi-level C-test confirm the main findings by Babaii and Moghaddam [4], i.e., text characteristics at different levels have potential effects on the C-test performing strategies.

We acknowledge unavoidable restrictions such as a limited sample that did not allow conducting standard validity tests. However, given the qualitative approach adopted, this was not the primary purpose of the study.

Further studies of TAP in regard to the C-test may also explore the link between the failure frequency with the text difficulty level, to compare the current finding with Salehi and Sanjareh [39], and Babaii and Fatahi-Majd [29]. In addition, academic version versus general version of international English as a Second/Foreign Language could be explored to help administer test-preparation and test-taking. This may also involve both topic types and test patterns.

Ultimately, following the results, pedagogical recommendations could be developed to build individual trajectories for reading test-taking and/or learning vocabulary. Teaching implications of this study also include considering text selection for assessment [4] and/or preparing redundancy tests for assessment at different proficiency levels [40]. It is the introspective data of the reading ability of participants that give teachers the exact estimation of candidates' abilities to identify the source of problem and use interdependent strategies in performance. The findings might be of interest for assessors as well.

References

1. Klein-Braley, C. & Raatz, U. (1984) A survey of research on the C-Test 1. *Language Testing*. 1 (2). pp. 134–146.

2. Dörnyei, Z. & Katona, L. (1992) Validation of the C-test amongst Hungarian EFL learners. *Language Testing*. 9 (2). pp. 187–206.

3. Klein-Braley, C. (1997) C-Tests in the context of reduced redundancy testing: an appraisal. *Language Testing*. 14 (1). pp. 47–84.

4. Babaii, E. & Moghaddam, M.J. (2006) On the interplay between test task difficulty and macro-level processing in the C-test. *System*. 34 (4). pp. 586–600.

5. Sarapuu, I. & Alas, E. (2016) Developing a c-test to measure language ability as an alternative to a skills-based test. *Eesti Rakenduslingvistika Ühingu Aastaraamat*. 12. pp. 237–252.

6. Hiser, E.A. & Ho, K.S.T. (2016) C-Tests in Vietnam: An Exploratory Study of English Proficiency. *Electronic Journal of Foreign Language Teaching*. 13 (2). pp. 184–202

7. Eckes, T. & Grotjahn, R. (2006) A closer look at the construct validity of C-tests. *Language Testing*. 23 (3). pp. 290–325.

8. Phakiti, A. (2003) A closer look at the relationship of cognitive and metacognitive strategy use to EFL reading achievement test performance. *Language Testing*. 20 (1). pp. 26–56.

9. Babaii, E. & Ansary, H. (2001) The C-test: a valid operationalization of reduced redundancy principle. *System*. 29 (2). pp. 209–219.

10. Daller, M. & Xue, H. (2007) Lexical richness and the oral proficiency of Chinese EFL students. In: Daller, H., Milton, J. & Treffers-Daller, J. (eds) *Modelling and Assessing Vocabulary Knowledge*. Cambridge: Cambridge University Press. pp. 150–164.

11. Drackert, A. & Timukova, A. (2019) What does the analysis of C-test gaps tell us about the construct of a C-test? A comparison of foreign and heritage language learners' performance. *Language Testing*. 37 (1). pp. 107–132.

12. Mashad, I. (2008) Another look at the C-Test: A validation study with Iranian EFL learners. *The Asian EFL Journal Quarterly*. 10 (1). pp. 154–180.

13. Jafarpur, A. (1995) Is C-testing superior to cloze? *Language Testing*. 12 (2). pp. 194–216.

14. Hastings, A. (2002) In defense of C-testing. In: Grotjahn, R. (ed.), *Der C-Test. Theoretische Grundlagen und praktische Anwendungen*. Vol. 4. Bochum: AKS-Verlag. pp. 11–29.

15. Müller, A. & Daller, M. (2019) Predicting international students' clinical and academic grades using two language tests (IELTS and C-test): A correlational research study. *Nurse Education Today*. 72. pp. 6–11.

16. Prediger, S. et al. (2018) Language proficiency and mathematics achievement. *Journal für Mathematik-Didaktik*. 39 (1). pp. 1–26.

17. Daller, M.H. et al. (2011) Language dominance in Turkish-German bilinguals: methodological aspects of measurements in structurally different languages. *International Journal of Bilingualism*. 15 (2). pp. 215–236.

18. Lee, J.U., Schwan, E. & Meyer, C.M. (2019) *Manipulating the Difficulty of C-Tests*. arXiv preprint arXiv:1906.06905.

19. Wilmes, C. (2007) *Validation of a German language placement test based on a modified C-test procedure*. Doctoral Dissertation. University of Illinois at Urbana-Champaign.

20. Kaivanpanah, S. & Soltani Moghaddam, M. (2012) Knowledge sources in EFL learners' lexical inferencing across reading proficiency levels. *RELC Journal*. 43 (3). pp. 373–391.

21. Peretz, A.S. & Shoham, M. (1990) Testing reading comprehension in LSP: Does topic familiarity affect assessed difficulty and actual performance. *Reading in a Foreign Language*. 7 (1). pp. 447–455.

22. Zhang, L.J. (2001) Awareness in reading: EFL students' metacognitive knowledge of reading strategies in an acquisition-poor environment. *Language Awareness*. 10 (4). pp. 268–288.

23. Zhang, L.J. (2009) Chinese senior high school EFL students' metacognitive awareness and reading-strategy use. *Reading in a Foreign Language*. 21 (1). pp. 37–59.

24. Bengeleil, N. & Paribakht, T. (2004) L2 reading proficiency and lexical inferencing by university EFL learners. *Canadian Modern Language Review*. 61 (2). pp. 225–250.

25. Lee, J. W. & Schallert, D. L. (1997) The relative contribution of L2 language proficiency and L1 reading ability to L2 reading performance: A test of the threshold hypothesis in an EFL context. *TESOL Quarterly*. 31 (4). pp. 713–739.

26. Soemer, A. & Schiefele, U. (2019) Text difficulty, topic interest, and mind wandering during reading. *Learning and Instruction*. 61. pp. 12–22.

27. Matsumoto, K. (1993) Verbal-report data and introspective methods in second language research: State of the art. *RELC Journal*. 24 (1). pp. 32–60.

28. Rahimi, M. & Saadat, M. (2005) A verbal protocol analysis of a C-test. *International Journal of Applied Linguistics*. 8 (2). pp. 55–86.

29. Babaii, E. & Fatahi-Majd, M. (2014) Failed restorations in the C-test: Types, sources, and implications for C-test processing. In: Grotjahn, R. (ed.) *Der C-Test: Aktuelle Tendenzen/The C-Test: Current trends*. Peter Lang GmbH, Internationaler Verlag der Wissenschaften. pp. 263–276.

30. Henning, G. (1987) *A guide to language testing*. Rowley, MA: Newbury House Publishers.

31. Baumann, J.F., Jones, L.A. & Seifert-Kessell, N. (1993) Using think aloud to enhance children's comprehension monitoring abilities. *The Reading Teacher*. 47 (3). pp. 184–193.

32. Cohen J. (1988) Statistical power and analysis for the behavioral sciences. 2nd ed. Hisdale, NJ: Lawrence Erlbaum Associates.

33. Cohen, A.D. (2014) *Strategies in learning and using a second language*. 2nd ed. London and New York: Routledge.

34. Iwahori, Y. (2008) Developing Reading Fluency: A Study of Extensive Reading in EFL. *Reading in a Foreign Language*. 20 (1). pp. 70–91.

35. Ryan, K.E. & Bachman, L.F. (1992) Differential item functioning on two tests of EFL proficiency. *Language Testing*. 9 (1). pp. 12–29.

36. Tsai, Y.C. & Li, Y.C. (2012) Test anxiety and foreign language reading anxiety in a reading-proficiency test. *Journal of Social Sciences*. 8 (1). pp. 95–103.

37. Liu, M. (2006) Anxiety in Chinese EFL students at different proficiency levels. *System*. 34 (3). pp. 301–316.

38. Busch, D. (1982) Introversion-extraversion and the EFL proficiency of Japanese students. *Language Learning*. 32 (1). pp. 109–132.

39. Salehi, M. & Sanjareh, H.B. (2013) On the comparability of C-test and Cloze: A verbal protocol approach. *English for Specific Purposes World*. 14 (3). pp. 1–17.

40. Ajideh, P., Farrokhi, F. & Nourdad, N. (2012) Dynamic Assessment of EFL Reading: Revealing Hidden Aspects at Different Proficiency Levels. *World Journal of Education*. 2 (4). pp. 102–111.

Appendix 1

Directions.

Перед вами небольшой текст, в котором необходимо заполнить пропуски. Первое и последнее предложения оставлены в неизменном виде, а в оставшихся предложениях в каждом втором слове удалена часть слова – если это слово с четным количеством букв, то удалена половина слова, если это слово с нечетным количеством букв, то удалена большая часть букв. Длина пропуска в тексте не зависит от количества удаленных букв.

Восстановите слова с пропусками так, чтобы предложения и текст имели смысл. Выполняя задания, рассуждайте вслух и озвучивайте все ваши мысли и гипотезы, какими бы абсурдными они вам ни казались.

1. Rats

According to French scientists, rats are much cleverer than we thought. Attempts t_____ reduce t_____ rat popul_____ have consis_____failed, a_____ environmental exp_____ have expl_____ this fa_____ by clai_____ that ra_____ have deve_____ a resis_____ to poi_____. According t_____ scientists w_____ met i_____ Lyons t_____ discuss rese_____ into r_____ intelligence, i_____ now se_____ that t_____ rat i_____ clever eno_____ to outs_____ humans. Rats have an ability to communicate their thoughts.

2. Getting a Return From Training

The employees know their jobs. Th_____ have a_____ successfully comp_____ initial trai_____ courses a_____ they ha_____ had ple_____ of dir_____ daily exper_____. Nonetheless, comp_____ instinctively fe_____ they co_____ get mo_____ of th_____ staff b_____ giving th_____ further trai_____ in ar_____ related t_____ their wo_____. When i_____ comes t_____ investing i_____ such st_____ training, how_____, there i_____ a fundamental par_____.Training remains an absolutely luxury item.

3. International Law

The term “national law” is used to mean the internal legal rules of a particular country, in contrast to international law, which deals with the external relationships of a state with other states. There i_____ no wo_____ government o_____ legislature iss_____ and enfo_____ laws t_____ which a_____ nations a_____ subject. T_____ international le_____ order h_____ no sin_____ governing bo_____ and oper_____ by agre_____ between sta_____. This me_____ that t_____ creation, interpr_____ and enfor_____ of intern_____ law l_____ primarily i_____ the ha_____ of sta_____ themselves. Its scope and effectiveness depend on the sense of mutual benefit and obligation involved in adhering to the rules.

Использование методики Think-Aloud Protocols для выявления стратегий выполнения многоуровневого С-теста

Бабаии Э., Пермякова Т.М., Поздеева Е.В.

Ключевые слова: С-тест, когнитивные стратегии, оценка языковых компетенций, методика «думай вслух»

DOI: 10.17223/19986645/73/1

В данном исследовании рассматривается процесс выполнения многоуровневого С-теста на английском языке. С-тест представляет собой набор коротких текстов, в которых отсутствует часть каждого второго слова. Учащимся необходимо заполнить пропуски, т.е. восстановить тексты до исходного состояния. С-тесты широко используются как надежный инструмент для измерения уровня владения иностранным языком.

При проведении исследования авторы ставили перед собой две цели: 1) определить когнитивные стратегии, которыми пользуются испытуемые-носители русского языка при выполнении многоуровневого С-теста на английском языке; 2) выявить соотношение между частотой использования когнитивных стратегий и уровнем сложности С-теста. В эксперименте приняли участие 20 испытуемых, уровень языковой компетенции каждого из которых был определен как B2 (согласно шкале общеевропейских компетенций владения иностранным языком). Испытуемые выполняли С-тест, состоящий из трех частей, при этом уровень сложности каждой последующей части возрастал от B1 до C2. Для того чтобы определить, какими когнитивными стратегиями испытуемые пользовались при выполнении теста, авторы использовали методику Think-Aloud Protocols или метод «думай вслух».

В результате исследования были выявлены семь когнитивных стратегий, применяемых при выполнении С-теста: «прогнозирование», «использование умозаключений», «стратегия вопроса», «сопоставление / установление связей», «перечитывание / использование подсказок», «выявление проблемы» и «рефлексия». Было установлено, что частота использования различных стратегий меняется в зависимости от уровня сложности текста. Так, например, при выполнении теста уровня B1 наиболее часто применяемыми стратегиями были «использование умозаключений» и «перечитывание / использование подсказок», в то время как при выполнении теста уровня B2 возрастала частота использования таких стратегий, как «прогнозирование», «сопоставление / установление связей» и «рефлексия». При выполнении теста уровня C2 наиболее активно используемыми стратегиями были «стратегия вопроса» и «выявление проблемы». Проверка с помощью критерия хи-квадрата позволила сделать вывод, что различия в частоте использования когнитивных стратегий являются статистически значимыми и уровень сложности текста может инициировать использование различных подходов к выполнению С-теста.

Полученные результаты могут оказаться полезными для дальнейших исследований стратегий, применяемых при выполнении тестов, в частности для исследований, проводимых с участием испытуемых-носителей разных языков. Кроме того, полученные результаты могут быть использованы для разработки подходов к изучению иноязычной лексики, обучению чтению и выполнению тестов на иностранном языке, как общего характера, так и с учетом индивидуальных особенностей учащихся.