
M. Medved Krajnović & R. Šamo, *Testing Foreign Language Proficiency...SRAZ LII*, 283-299 (2007)

UDC 371.279.7:81'243

Original scientific paper

Received on 5 September 2007

Accepted for publication on 3 April 2008

Testing Foreign Language Proficiency – What Verbal Protocols Reveal

Marta Medved Krajnović

Faculty of Humanities and Social Sciences, Zagreb

Renata Šamo

Juraj Dobrila University of Pula, Pula

The paper deals with several issues related to testing second language learners'¹ communicative language competence. It concentrates on verbal protocols as a means of gaining insight into learners' test taking strategies and cognitive processes, and as one of the ways of checking test validity and reliability. The research part of the paper presents results of the application of the verbal protocol technique with primary school children who were faced with three different cloze-test tasks. The research proved that verbal protocols can be used with learners of this age group, and that the application of verbal protocols is very important for a more reliable interpretation of learners' test results.

1. The notion of communicative language competence or language proficiency²

The process of formal teaching and learning is inextricably linked to the process of testing what has been taught or learned. And what is usually

¹ The term 'second language learner' is used as a cover term to imply both second and foreign language learners.

² The terms of communicative competence, communicative language competence and language proficiency will be used interchangeably since for the purpose of this paper their distinction (if existing at all) is not relevant.

taught or learned are pieces of information, knowledge about the world, knowledge about how to do something.

When learning a foreign language, we learn a certain degree of 'factual' information – elements of language such as vocabulary, grammar, as well as metalinguistic information about these elements. Through the new language we learn, we also learn about cultural and social elements related to it, i.e. we acquire new concepts, new knowledge of the world, and we learn how to use this knowledge in communicative situations, e.g., how to listen effectively, how to use different reading strategies, how to write different types of texts, how to speak in formal and informal contexts.

All these 'learnable' elements that we have just mentioned could be considered part of communicative language competence. For some 30 years there has been extensive (and inconclusive) research and theorising about definitions and models of communicative language competence, the key question being whether communicative competence is one unitary thing or a multicomponential construct, and if the latter, what its components are and how they are inter-related. There has also been an ongoing discussion (for a comprehensive overview see Bagarić 2007) about the relationship of underlying competence and actual performance. The main concern was whether performance reflects competence, or one can only vaguely infer about competence on the basis of performance, since performance is affected by many inside (e.g., processing, affective) and outside factors (e.g., context of language use, interlocutors). And on the top of all these questions that were equally shared by second language acquisition researchers and researchers specializing in the field of language testing, the latter focused on one more crucial question: How much and what part of communicative language competence, or language proficiency, is reflected through performance in such artificial second language use situation as language tests are?

It is evident that in order to answer any of the above questions, the first thing that has to be at least partially explained is the construct of communicative competence. And despite the already mentioned inconclusiveness of explanations, we can say that today there are some landmarks, some models and 'facts' about communicative language competence that have been accepted and used (e.g., for the purpose of teaching or testing) more than others.

The researchers nowadays agree that communicative competence is a multidimensional construct consisting of several components. The number and name of these components varies from model to model (e.g., Canale 1983; Savignon 1983; Bachman 1990; Bialystok 1994; Celce-Murcia, Dörnyei and Thurrell 1995; Bachman and Palmer 1996; Council of Europe 2000), but if we wanted to single out the most frequent components and give them more or less transparent names, they would be the following: linguistic competence, pragmatic/sociolinguistic competence and strategic competence. These could be considered as higher order components and in different models they have different subcomponents. Researchers do not agree on the type and number of these subcomponents either, not to mention their relationship, but it seems that there is a growing consensus that communicative language competence is an immensely complex construct, that not all (sub)components have equal weight, and that probably their structure and relationship is partly shaped by different experiences second language users had while learning/acquiring a second language.

2. Testing foreign language proficiency

Nowadays when knowing foreign languages has become a prerequisite of a successful professional and private life, reliable and valid testing of foreign language proficiency has become a must. And as already said, in order to devise an appropriate assessment or testing procedure, we should know what the construct of communicative language competence is. In the history of language testing three different periods, based on rather different views of the construct of communicative competence, could be singled out. We can call these periods/views 'structuralist' (which dominated in the 1960s), 'unitarist' (dominated during the 1970s and early 1980s) and 'modernist' (from late 1980s onwards). During the first period, the predominant view was that language has a firm structure consisting of the four skills (listening, speaking, reading and writing) and their components (vocabulary, grammar, phonology/orthography, rate/fluency), and that language tests measuring each of these components separately could and should be devised (e.g., Harris 1969). Tests that dominated in

this period were the so-called discrete points tests (e.g., multiple choice grammar test). As a reaction to this fractionist view of language proficiency came a radically opposite view whose main proponent was Oller (1979). Oller and his followers claimed that language proficiency is one unitary thing and that only tests aiming at measuring this unitary factor (the so-called g-factor) were the appropriate ones. This was the era of integrative-skills tests when, for example, cloze-tests were devised and most often used. As the research into the phenomena of communicative language competence and testing issues accumulated and matured, the new, 'modern' era of language testing came into being. The landmarks during this period were the works of Bachman (1990, 2004), Alderson, Clapham and Wall (1995), Bachman and Palmer (1996), McNamara (2000, 2001a, 2001b), Shohamy (2001), Hughes (2003), McKay (2006). During this era researchers gradually became aware of the limitations of the assessment process stemming from the already mentioned complexity of the construct of communicative competence, non-straightforward relationship between competence and performance, artificiality of testing situations (despite test designers' striving for authenticity). They accepted these limitations and, with a positive belief that future research will bring more answers, have been trying to devise testing procedures which are valid, reliable and appropriate to a particular context of testing. In addition to this, procedures have been developed in order to investigate not only the testing product (e.g., learners' answers), but also the underlying cognitive processes or strategies learners use while solving a language problem. The next section will deal with this.

3. The role of verbal protocols in language testing

As said above, almost every model of communicative language competence that has had a wider influence and application in the field of language testing includes strategic competence as a component that has to be taken into consideration while evaluating language users' proficiency. Research into strategies that language learners/users employ while learning or using a second language was extremely prolific during the late 1980-ies and early 1990-ies (Faerch and Kasper 1983; Bialystok 1990; O'Malley and Chamot 1990; Oxford 1990, 1996; Wenden 1991; McDonough

1995; Cohen 1998). A part of this research also included investigation into learners' test taking strategies, and even early results showed that one can be quite mistaken about learners' language proficiency if explaining it only on the basis of learners' test results (for an overview of this early research see Medved Krajnović 1999). Test taking strategy research proved to be very useful in the process of test development and more reliable assessment of learners' communicative language proficiency.

However, in order to investigate learners' test-taking behaviour, new research methods and techniques had to be developed. Since learners themselves were the only ones who could tell researchers about the processes and strategies used in a particular task, they were asked to do so either in written or spoken form. In this paper we will concentrate on the spoken form of these reports, the so-called verbal protocols.

Although, in order to investigate problem-solving and memory skills, protocol analysis was used in the field of psychology already in the first half of the 20th century, it has been accepted by second language acquisition researchers only in the past two to three decades (see Afflerbach 2001; Anderson & Vandergrift 1996; Cohen 1998; O'Malley and Chamot 1990; Wenden 1991). According to Jourdenais (2001), protocol reports can be basically grouped into: *introspective*, *retrospective*, and *think-alouds*. What differentiates them is the type of information they elicit and time when they are collected. In introspective reports processing strategies are explained simultaneously with task performance. In contrast, retrospective reports are collected upon task completion as performers are asked to think back upon and report the processes they applied and thoughts they had during the task. Just like the first mentioned group of reports, think-alouds are also collected during task execution, but they do not ask test-takers to interpret their cognitive behaviour. Test-takers are asked to 'think aloud' while performing the task, verbalising the procedure from the moment when it starts until it finishes. It is generally assumed that information most recently attended by the task performer is available for verbal report. According to memory and information processing models (see Ellis 2001; Ericsson & Simon 1984), this accessible information comprises the contents of working or short-term memory, where information is first processed and stored to be later possibly

incorporated into long-term memory. Being collected simultaneously with task performance, introspective and think-aloud reports are designed to probe directly into working memory. Retrospective reports, however, may collect information from a performer's short-term memory (accessible for a very limited time, though), but they may also require the retrieval of information from his/her long-term memory.

Verbal protocols have been increasingly used in second language acquisition (SLA) research as a means of gaining insight into second language users' linguistic, cognitive and emotional processes while performing a task in a second language. The use of such a method can help researchers in getting a better understanding of why test-takers respond the way they do, how they understand the task, how they feel during the performance. However, protocol use is not without its weaknesses. Problem issues that arise could be summarised as follows: whether or not test takers are able to accurately recall (because of memory and other constraints) the thoughts they have during task completion; whether or not they report what they feel researcher wants to hear rather than what they actually experience (protocol reliability); whether or not test takers are actually able to verbalise everything that they are thinking (protocol data completeness); whether or not they report certain behaviour but not other (elicitation techniques influence); whether or not second/foreign language users have access to the metalinguistic information, i.e. have sufficiently developed metalinguistic ability to describe their behaviour (see Cohen, 1998; O'Malley and Chamot 1990).

Fortunately, a number of concerns regarding the use of protocol data collection and analysis can be diminished, and even overcome, by a careful research design (including training sessions for potential candidates), elicitation of verbalisations, coding of data. Another way to almost entirely eliminate some problems is to gather different types of protocol data. For example, collecting think-alouds, in which verbalisations are provided simultaneously with task performance, will ease concerns about memory constrains. Besides, they do not require test-takers' interpretation or explanation of behaviour, so they may provide researchers with a more comprehensive and accurate picture of L2 user's on-line processing. However, we should note that some candidates, despite training sessions,

may not be able to provide detailed information about their actual performance, and others may simply be very uncomfortable with such a task. In such cases, additional retrospective reports may provide useful additional data.

4. Aim of the study

The aim of the present study, which was carried out as part of a larger scale research into second language reading strategies (Šamo, 2006), was to see what verbal protocols can reveal about young second language users' behaviour during a very specific second language task – a cloze-test. We have already mentioned that cloze-tests dominated the field of language testing during the era when language proficiency was considered to be one unitary thing. However, there have been many controversies related to this testing technique, and the two basic are:

- . the idea of communicative language competence as one unitary thing was refuted;

- . research into the reliability and validity of cloze-tests showed that even educated native speakers scored differently on these tests, depending on the type of text and, particularly, on the words being deleted (e.g., if content words were deleted, even native speakers sometimes had problems in retrieving the exact word).

However, since the aim of our study was to 'test' verbal protocols as a research method, we thought that a controversial cloze-test would be a very appropriate elicitation technique.

5. Sample, instruments and procedure

The study included 37 Croatian primary school learners of English as a foreign language aged 13-14 years. The age when they had started learning English ranged from six to ten years. At the time of testing, English was a compulsory school subject for all participants.

Data relevant for our study were elicited by three different cloze-tests of a classical type (i.e. every n-th word omitted, first and last sentences

intact) and two types of verbal protocols - think-alouds and retrospective reports. Cloze-tests consisted of 22, 32 and 38 gaps respectively, the completion of which required both candidates' linguistic knowledge (e.g., for filling in grammar related gaps) and their knowledge of the world around them (e.g., the topics of cloze passages were traffic, holidays, music). The cloze-tests were individually administered to participants over a four-month period, and the gap-filling procedure was simultaneous with the think-aloud procedure. Interviews, i.e. retrospective reports, followed immediately afterwards. The entire procedure was audio-taped, transcripts coded and analysed qualitatively and quantitatively. Prior to the research, the participants had been trained on how to verbalise their thoughts while doing the task. Although all of them had been eager to participate in the research, their motivation slightly decreased as the training progressed. Some participants considered each cloze-task more demanding than the previous one and additionally burdened by the think-aloud procedure. It was not easy for them to verbalise their thoughts, and often when asked why they thought a particular solution was appropriate, they just shrugged and answered that this was the way to be, or that to them it simply seemed the best option. Yet, they did try hard to find as many solutions as possible in order to do the task and, in the end, stated that they enjoyed the study, viewing it mainly as a welcome change and challenge in their everyday school life.

6. Results and discussion

The study resulted in 111 verbal protocols and we performed both qualitative and quantitative analyses on the obtained data. In Table 1 we present test-taking problem issues that the participants reported during think-alouds and interviews. It is interesting to observe that the interview data only confirmed the problems participants already mentioned in think-alouds.

Table 1: Frequency and percentage of students' problem issues

Task No.	Comprehension problems				Task-related problems						Knowledge and/or self-confidence problems			
	Word level		Text (part) level		Task format		Task demands		Physical terms of procedure		Lack of knowledge		Lack of security	
	F	%	F	%	F	%	F	%	F	%	F	%	F	%
Task 1	26	70.3	10	27.0	6	16.2	8	21.6	8	21.6	9	24.3	6	16.2
Task 2	6	16.2	7	18.9	1	2.7	3	8.1	3	8.1	6	16.2	1	2.7
Task 3	4	10.8	7	18.9	-	-	3	8.1	2	5.4	4	10.8	2	5.4

As shown in Table 1, learners' problems could be grouped into three main areas. Firstly, there were comprehension problems at the word level as well as at the text level (the entire text or a paragraph); secondly, there were task-related problems associated with participants' inexperience with a cloze-test as a task format, with task requirements which asked for complete and accurate text reconstruction by adding the most appropriate missing words, and with the circumstances in which the research procedure was carried away (unusual for participants because of individual audio-recording in a separate room, simultaneous test-taking and think-aloud procedures, and the presence of the researcher and her assistant); the third set of problems was related to the participants' general sense of having inadequate linguistic and background knowledge and/or self-confidence in such research situation.

The first task (cloze-test 1) was the only one where the test takers experienced serious comprehension problems at the word recognition level, probably mainly due to the fact that the participants were later allowed to ask for help when faced with an unfamiliar word, or a situation when they simply could not remember a word. We noticed that attempts at guessing some words had been time-consuming, demanding and eventually fruitless, so we did not want participants to lose motivation for solving other types of comprehension problems, i.e. at the higher, text-processing level. However, we generally guided them towards making

their own conclusions about meaning, instead of offering them ready-made solutions at once.

From the above-presented table we can also see that comprehension problems above the word recognition level were more obvious in task 1 than in tasks 2 and 3. We suppose that the main reason for this was in the narrative scheme of the first text, which required very careful reading and identifying the roles that each character had in the story – not an easy task when the two key words naming the main characters were among the most difficult to guess and remember. Even after guessing correctly what the text was about, many participants showed some reluctance to stick to their first interpretation just because they could not understand what ‘tortoise and snail’ meant, or how to view them in the given plot.

When compared with the second and third task, for the first one the learners also reported the highest percentage of task-related problems, which probably can be best explained by the fact that participants found themselves in this specific research situation for the first time. However, the influence of a cloze-test format significantly decreased in the second, and completely disappeared in the third task, while the two other task level problems remained until the end of the research procedure, but with the low percentage of occurrence. It is reasonable to suppose that, as the whole procedure advanced, the participants got accustomed to the research context and their role in it. Besides, a relationship of mutual trust between them and the researcher developed. According to some of their comments, they started viewing the task as an anxiety-free co-operation and a stimulating experience.

Unsurprisingly, during the first task the participants also reported lack of knowledge and self-confidence problems more frequently than in the rest of the procedure.

What follows are the excerpts from qualitative data supporting the above mentioned, i.e. highlighting the main problems learners felt they faced, especially at the beginning stages of the research. For cloze-test 1 they reported the following problems:

- being unable to recognise some words presented in the text, e.g., ‘...

I had no idea that these words existed, so I had to think about what they might mean ...’ (VP 2/1), ‘... it’s difficult as I don’t know so many words.’ (VP 8/1), ‘

This is my problem - I can't understand some words. (VP 12/1) or '... the words are too difficult for me ...' (VP 14) as well as being unable to remember some words to be added to the text, e.g., '... I can't remember the sentence, not the sentence but the word, to write down here ...' (VP 9/1), '... I can't connect it to any of these words - can't remember a single word.' (VP 16/1)

- being unable to understand the text completely, not only because they could not recognise the key words, but also because of the unexpected narrative scheme - description of an event unusual in animals' life (tortoise and snail as the main characters in a car accident), e.g., *'I didn't quite know if this was about animals or people?!'* (VP 28/1)

- being confused by the demands of a specific, previously unknown task (cloze-test): e.g., *'Because I've never done this before, this test with missing words, and I know many English words, so it's such a mess in my head.'* (VP 35/1)

- feeling that they require better competence in English and more confidence at suggesting solutions, e.g., *'... maybe, I should know this but I have no idea ...' (VP 14/1), '... it was all ... it seemed to me that anything could be good ...' (VP 25/1), 'There are some words that I might have written down in a different way, not sure (VP 5/1).'*, *'Not quite sure, for example, nouns and ...' (VP 19/1) .*

The qualitative analysis of the data obtained during the second gap-filling task (cloze-test 2) showed problems very similar to the ones in the first task, and some additional ones such as:

- being unable to understand a part of the text, e.g., *'... At the beginning I did understand something but this end of the text is not so easy for me.'* (VP 4/2), *'This part confuses me, I can't understand it all.'* (VP 4/2)

- text not being appropriate to their age and life experiences (a student who goes camping in mountains and a businessman who goes cruising somewhere in the Caribbean as the main characters), e.g., *'... this one is a bit more difficult ... because of its plot ...' (VP 34/2).*

Finally, data obtained in the third stage indicated no serious problems, as we have previously presented (see Table 1). Therefore, this time we summarise the main advantages participants ascribed to the third text

(cloze-test 3) in order to explain why it was found easier by the majority of them:

•text being more appropriate for teenagers (two young people attending a rock concert) , e.g., ' ... it's about music, not boring like the first one, the second one was also good but I like this one more ...' (VP 20/3), ' ... as the topic is interesting to me, it gives me more potential to do the test better.' (VP 29/3); consequently the text being less demanding in terms of comprehension, e.g., ' ... I could understand more ... what it is about ...' (VP 3/3), 'Since the very beginning I could understand it and got absorbed in it, I immediately knew what it was about.' (VP 4/3), '... more familiar words to connect sentences better than last time ...' (VP 31/3); therefore, the whole task being experienced as not so demanding, e.g., '... this one is the easiest task (...) the words not so complicated, easier to understand what to put where ...' VP 23/3), '... maybe a bit easier than the previous ones - because of the words that we had to insert ...' (VP 28/3).

What is important to observe is that the participants' impressions about the amount of problems they faced in each of the three tasks correlate well with the results they obtained on these tasks. These results are presented in Table 2:

Table 2: Summary statistics for students' performance on cloze-tests

	Cloze-test 1	Cloze-test 2	Cloze-test 3
M	19.73	29.30	38.73
SD	11.43	14.99	17.96
Min score	0	0	0
Max score	38	55	65
Max possible score	44	64	76
% success	44.8 %	46 %	51 %

Table 2 presents the third reading task as the most successful one, and we suppose it is because of the following:

. the topic of the text (rock-concert) that made our participants perceive it as the closest to their life experiences - they could easily identify with

the main characters and therefore felt more secure and motivated to do the task;

. the language of the text - there were just a few words not quite known to the participants, and they could recognise them on their own or guess the meaning from the given context;

. better familiarity with the task type than at the beginning of the study and, therefore, a higher degree of participants' self-confidence in the later stages of the research;

. more relaxing atmosphere characteristic of this research stage.

Therefore, we could conclude that in order for research to be satisfactorily carried out, participants should be prepared for what will be expected from them, i.e. have a pre-testing training (in our case on cloze-testing and think-aloud procedure), and they should be allowed time to gather some real, on-task experience and establish a relationship with the researcher.

7. Concluding remarks

We believe that our data point to the usefulness of the verbal protocol procedure. The verbal protocols revealed some of the learners' test-taking strategies (e.g., guessing, relying on language intuition), but they primarily showed what else in addition to linguistic (for our participants mostly lexical) knowledge can influence candidates' results on a cloze-test. We would like to suggest that results on cloze-tests are influenced by test-takers' familiarity with the task format, i.e. testing technique, and familiarity and appropriateness of the topic of the text. Being familiar with the task and text topic can boost participants' self-confidence, which then positively affects performance. Our data also showed how important it is for second language learners that the key words in a reading passage are not too unfamiliar, or used in an unusual context.

And although the just mentioned issues are already well known to testing experts, we think that special value of our research lies in the fact that we have shown that even primary school children are aware of these issues, and that despite children's feelings of uneasiness and difficulty

related to verbal protocols, they all managed to follow the requested procedure. This is important to realise since verbal protocols have been mostly used with adults. Data from children's test taking verbal protocols can reveal some specificities of testing foreign language proficiency of primary school learners, and consequently help us in designing better tests for this age group.

REFERENCES:

- Afflerbach, Peter. 2001. Verbal Reports and Protocol Analysis, *Handbook of Reading Research* [ed. by Pearson, David, Kamil, Michael, Mosenthal, Peter and Barr, Rebecca.], New York: Longman, pp. 163-179.
- Alderson, Charles, Clapham, Caroline and Wall, Dianne. 1995. *Language Test Construction and Evaluation*, Cambridge: Cambridge University Press.
- Anderson, J. Neil and Vandergrift, Laurens. 1996. Increasing metacognitive awareness in the L2 classroom by using think-aloud protocols and other verbal report formats, *Language learning strategies around the world: Cross-cultural perspectives*. [ed. by Oxford, L. Rebecca], Honolulu: University of Hawai'i, Second Language Teaching & Curriculum Center, pp. 3-18.
- Bachman, Lyle. 1990. *Fundamental Considerations in Language Testing*, Oxford: Oxford University Press.
- Bachman, Lyle. 2004. *Statistical Analyses for Language Assessment*, Cambridge: Cambridge University Press.
- Bachman, Lyle and Palmer, Adrian. 1996. *Language Testing in Practice: Designing and Developing Useful Language Tests*, Oxford: Oxford University Press.
- Bagarić, Vesna. 2007. *Struktura komunikacijske kompetencije u stranom jeziku*, [The Structure of Foreign Language Communicative Competence], Unpublished doctoral dissertation, Zagreb: Faculty of Philosophy.
- Bialystok, Ellen. 1990. *Communication Strategies*, Oxford: Blackwell.
- Bialystok, Ellen. 1994. Analysis and control in the development of second language proficiency, *Studies in Second Language Acquisition*, 16, pp. 157-168.

- Canale, Michael. 1983. From communicative competence to language pedagogy, *Language and Communication* [ed. by Richards, Jack and Schmidt, Richard], London: Longman, pp. 2-27.
- Celce-Murcia, Marianne, Dörnyei, Zoltán and Thurrell, Sarah. 1995. Communicative Competence: A Pedagogically Motivated Model with Content Specifications, *Issues in Applied Linguistics*, 6, pp. 5-35.
- Cohen, Andrew. 1998. *Strategies in Learning and Using a Second Language*, London and New York, Longman.
- Council of Europe. 2000. *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*, Cambridge: Cambridge University Press.
- Ellis, C. Nick. 2001. Memory for Language, Cognition and Second Language Instruction [ed. by Robinson, Peter], Cambridge: Cambridge University Press, pp. 33-68.
- Ericsson, K. Anders, Simon, A. Herbert. 1984. *Protocol Analysis: Verbal Reports as Data*, Verlag: MIT.
- Faerch, C., Kasper, Gabriele. 1983. *Strategies in Interlanguage Communication*, London: Longman.
- Harris, David. 1969. *Testing English as a Second Language*, New York: McGraw Hill.
- Hughes, Arthur. 2003. *Testing for Language Teachers*, revised ed., Cambridge: Cambridge University Press.
- Jourdenais, Renée. 2001. Cognition, instruction and protocol analysis, *Cognition and Second Language Instruction* [ed. by Robinson, Peter], Cambridge: Cambridge University Press, pp. 354-375.
- McDonough, Steven. 1995. *Strategy and Skill in Learning a Foreign Language*, London: Edward Arnold.
- McKay, Penny. 2006. *Assessing Young Language Learners*, Cambridge: Cambridge University Press
- McNamara, Tim. 2000. *Language Testing*, Oxford: Oxford University Press.
- McNamara, Tim. 2001a. Rethinking alternative assessment, *Language Testing*, 18, pp. 329-332.
- McNamara, Tim. 2001b. Language assessment as social practice: challenges for research, *Language Testing*, 18, pp. 333-349.

- Medved Krajnović, Marta. 1999. Učeničke strategije i testovi stranih jezika, *Strani jezici*, 28, pp. 149-155.
- Oller, John. 1979. *Language Tests at School: A Pragmatic Approach*. London: Longman.
- O'Malley, Michael, Chamot, Anna. 1990. *Learning Strategies in Second Language Acquisition*, Cambridge: Cambridge University Press.
- Oxford, Rebecca. 1990. *Language Learning Strategies: What every teacher should know*, Boston: Heinle & Heinle Publishers.
- Oxford, Rebecca. 1996. *Language Learning Strategies around the World: Cross-cultural Perspectives*, Hawai'i: University of Hawai'i Press.
- Savignon, Sandra. 1983. *Communicative Competence: Theory and Classroom Practice. Texts and Contexts in Second Language Learning*, Reading, Mass: Addison-Wesley Publishing Company.
- Shohamy, Elana. 2001. Democratic assessment as an alternative. *Language Testing*, 18, pp. 373-391.
- Šamo, Renata. 2006. *Analiza strategija čitanja uspješnih i manje uspješnih učenika engleskog kao stranog jezika*, [Analysis of L2 Reading Strategies Used by Good and Poor EFL Learners], Unpublished doctoral dissertation, Zagreb: Faculty of Philosophy.
- Wenden, Anita. 1991. *Learner Strategies for Learner Autonomy*, London: Prentice Hall International.

TESTIRANJE ZNANJA STRANOGA JEZIKA – ŠTO NAM OTKRIVAJU USMENA IZVJEŠĆA

U članku se raspravlja o problematici testiranja komunikacijske jezične kompetencije učenika stranoga jezika. Akcenat je na usmenim izvješćima učenika kao načinu da se stekne uvid u učeničke strategije rješavanja testova i kognitivne procese koje pri tom koriste, kao i o načinu provjere valjanosti i pouzdanosti testa. Prikazani su i rezultati istraživanja u kojem su učenici osnovne škole trebali primijeniti verbalne protokole dok su rješavali tri različita cloze-testa. Istraživanje je potvrdilo da su osnovnoškolci sposobni rješavati test i istovremeno izvještavati o tome te da su informacije dobivene tim usmenim izvješćima vrlo bitne za pouzdaniju interpretaciju učeničkih rezultata na testu.

Key words: testing, communicative language competence, second/foreign language learners, verbal protocols, test-taking strategies

Ključne riječi: vrednovanje, komunikacijska jezična kompetencija, učenik drugoga/stranoga jezika, usmena izvješća, strategije rješavanja testa

Marta Medved Krajnović
Department of English
Faculty of Humanities and Social Sciences, University of Zagreb
Ivana Lučića 3
10000 Zagreb, CROATIA
mmedved@ffzg.hr

Renata Šamo
Sveučilište Jurja Dobrile u Puli
Preradovićeva 1
52100 Pula, CROATIA
rsamo@inet.hr

