

ASSESSING GRAMMAR AND VOCABULARY - YES OR NO?

Jana Bérešová

Department of English Language and Literature, Faculty of Education, Trnava University in Trnava,
Priemyselná 4, P.O. BOX 9, 918 43 Trnava, Slovakia

Abstract

The paper discusses the current situation in language testing in some countries whose language curricula are influenced by the Common European Framework of Reference (CEFR), whose approach, however, to language testing does not precisely include the action-oriented approach promoted in the document of the Council of Europe. Various testing techniques used in different local contexts will be analysed and testing grammar and vocabulary (language in use) will call for possible solutions reflecting current trends fostering communicative competence – the ability to communicate effectively and spontaneously in real-life settings. Data achieved from the actual measurement of learners' performances will be shown and analysed.

Key words: *action-oriented approach, syntacto-centric approach, communicative approach*

1 INTRODUCTION

Teaching and testing are interrelated. Communicative language teaching has been followed by communicative testing, introduced by Bachman (1995) when he presented a theoretical model of communicative language ability, based on the concept of language as communication, recognizing the importance of the context, both discourse and sociolinguistic, in which language is used. In the last two decades, testing has developed so much that local tests started to compete with the international examinations and as O'Sullivan (2012) concludes the level of professionalism among local testers has begun to change, and there are many examples of excellent local examinations in which the knowledge of local culture, language and society contributes to the form and content of the test. This progress can be seen in many European countries, in which many examples of good tests can be found, mostly related to the secondary school-leaving examinations.

Another significant change emerged more than a decade ago, when the Common European Framework of Reference for Languages: Learning, teaching, assessment (CEFR) was officially published in 2001 and its purpose was oriented on learning, teaching and assessment. The CEFR started to attract teacher trainers, policy makers and test developers who could recognize a significant move towards communicative language competence, focusing on an action-oriented approach. This resulted in a number of calibrated descriptors for four communicative language skills whose acquisition might be measured to get a real picture of someone's target language competence.

The tendency to relate the local examinations to the CEFR became remarkably noticeable after the Manual was presented in 2009 by the Language Division of the Council of Europe. Local test developers were attracted to start the process of linking their tests to the CEFR and a great deal of effort has been done mostly in testing English in local context. English tests can be discussed from two perspectives. The first perspective is related to the countries where English has been taught for many years as language number 1 and is used in the country for communication as people are exposed to it through watching films in the original language, reading English books and the language became a natural part of their language portfolio. Another perspective is recognizable in the countries which started to focus on English in the 1990s and new conditions under which they started to develop their economies resulted in a significant prioritization of English as a lingua franca by the parents who were planning the future for their children. While western European countries regarded English as a

language of communication, the Eastern European countries were still approached by communicative language teaching, known more theoretically rather than applied practically at schools. As many features of traditional language teaching have survived, such as focusing on grammar, mostly grammatical structures different from their mother-tongue ones, the tests based on the communicative approach (context, authenticity, etc.) still include testing grammar and vocabulary.

2 COMMUNICATIVE LANGUAGE TESTING

The new approach to teaching has influenced the approach to testing. Being influenced by the model of communicative language ability, the test designers took a new look at the assessment of grammar, emphasizing that grammar cannot be treated as an isolated component of knowledge and must be assessed in the larger context of language in communication. Purpura states (2005) that the goal of every grammar test is to obtain information on how well a student knows or can use grammar to convey meaning in some situation where the target language is used. As noted by Purpura (2005, p. 4), "Currently, knowledge of grammar might be inferred from the ability to select a grammatically correct answer from several options on a multiple-choice test, to supply a grammatically accurate word or phrase in a paragraph or dialogue, to construct grammatically appropriate sentences, or to provide judgements regarding the grammaticality of an utterance". The analysis of some international tests has revealed that some of them (e.g. FCC) focus on tasks which show how well test-takers control their grammar and vocabulary, other measure knowledge of grammar inferring it from the ability to use grammar correctly while reading (e.g. TOEFL) or writing (e.g. City and Guilds).

Taking into account the primary goal of language learning – to foster communicative competence, or the ability to communicate effectively and spontaneously in real-life settings – the English test designers and item writers wanted to embrace a communication perspective of language. According to Purpura (2005) most linguists have embraced one of two general perspectives to describe linguistic phenomena. A syntactocentric perspective of language refers to observing and analyzing the way in which words are arranged in a sentence, while a communication perspective of language focuses on how language is used to convey meaning. While western European countries prefer assessing how language is used to convey meaning, the eastern European countries still prepare tests in which grammar and vocabulary are tested either included in the section related to testing reading or in a separate section referring to language in use. In both cases the tasks are based on context-based texts.

A communication-based perspective of language views grammar as a set of linguistic norms, preferences and expectations that an individual invokes to convey a host of pragmatic meanings, which are appropriate, acceptable and natural depending on the situation (Purpura 2005), which is preferably used in testing speaking and writing. The CEFR (2001) distinguishes between oral/written production and oral/written interaction. While range and accuracy might be considered qualitative aspects of language use, spoken language use needs fluency, interaction requires an interaction qualitative aspect and production calls for coherence. Cohesion theory focuses on how certain words link grammatical forms to meaning and contextual use. These cohesive ties may be used through grammatical cohesive devices on a sentential level.

Taking both perspectives into account, grammar can be tested communicatively with a concern for both form and meaning. Grammatical forms can be tested on closed-item or open-item tasks, and meaning is included in testing the ability to communicate either in writing or speaking.

Today's language tests are based on tasks simulating communication activities that the test-takers are likely to be engaged in real-life settings. Context-dependent testing assesses the test-takers' ability to

take account of contextual information in order to produce the expected response. Contextualisation means to what extent the test-takers are being assessed on the basis of their ability to engage with the context provided in the test (Read, 2001). School-leaving examination writing and speaking tasks predominantly focus on the appropriateness of students' vocabulary use in relation to the task, which means that the students are expected to meet the lexical requirements of the given situation.

3 TESTING GRAMMAR AND VOCABULARY IN SOME EUROPEAN COUNTRIES

Most of the countries of Central and Eastern Europe decided to change their approach to teaching and testing English after the 1990s having been influenced by the First Certificate in English (FCE). It is considered the most important of the Cambridge exams, officially claimed at upper-immediate level difficulty and consistent with CEFR B2 level. The FCE consists of five sections referring to reading, writing, use of English, listening and speaking. Use of English consists of four parts related to testing techniques such as multiple choice cloze, open cloze, word formation and key word transformation. While the first and third parts are related only to vocabulary testing, the second and the fourth focus on grammar. Open cloze is based on gap-filling requiring functional language. The key word transformation is based on grammatical structures.

In our research the focus was put on four countries: Croatia, Hungary, Slovakia and Slovenia. Apparently, the countries follow the samples of testing techniques from international tests, however, they call them differently and use different formats. While some countries prefer using cloze tests, others test by checking the accuracy of the text provided, using ticks for correct answers and crosses for the incorrect ones, which are expected to be rewritten correctly. Prevalingly, the tests are context-based and students are expected to deal with grammar and vocabulary in context.

3.1 Croatia

The test consists of a reading, listening and writing sections. The reading section comprises five tasks, out of which the last one refers to testing mostly vocabulary based on mostly content (notional) words. The choice of words is based on their specificity, so there are no other options, however, this type of the task may be quite demanding for item writers. Writing is based on one task, in which the students are expected to express their opinion on the given topic.

3.2 Hungary

As far as the school-leaving examination in English is concerned, it is delivered in two levels: intermediate and advanced. The latter consists of a reading comprehension section, four tasks testing grammar and vocabulary in a language in use section, and three tasks for assessing listening. The first task of language in use refers to word formation, the second one is multiple choice (based mostly on the selection of the proper grammatical structure), the third task measures the students' ability to use functional words, the fourth one is based on checking the line's correctness, ticking a correct answer and or deleting the word which does not fit in with the sense of the text. The communication-based perspective in testing grammar and vocabulary is applied in writing – two pieces of writing: writing an email and writing a letter giving an opinion on the given issue/issues.

3.3 Slovenia

The Slovenian test consists of three sections, the first one related to reading comprehension and language in use, the second one referring to listening comprehension, with the last one focusing on writing. Three tasks based on testing grammar and vocabulary make up two gap-filling testing techniques (writing the missing words and writing the correct forms of verbs), vocabulary is assessed

through word formation. The communication-based approach is recognizable in the writing section where the students are expected to write two pieces of writing, one essay the topic of which is selected by the student and one piece of writing referring to comments and opinions related to the question that is based on their reading experience during their studies.

3.4 Slovakia

The higher level test includes four sections: listening, language in use, reading and writing. All the sections are weighed equally, although the number of the items in the language in use section is 40. The first task is based on assessing grammar and vocabulary in a multiple-choice task (20 items), word formation covers 10 items and cloze test based on gap-filling of the functional words measures the ability students of syntactical relations of the words in an utterance.

The similarities and differences in the countries' approach to testing grammar and vocabulary is clearly seen from Table 1. The most frequent testing technique is word formation, used in three countries, while a cloze test, or gap filling task, which are based on the same principle, are used in three countries as well. Slovenia measures the ability to use proper tenses and verb forms in one task. On the other hand, Slovakia used to test this ability in the past, now cloze tests are more preferable as the focus is placed on the students' ability to read sentences thoroughly and recognize how the words are arranged in the sentence.

	Croatia	Hungary	Slovakia	Slovenia
Part 1		word formation	multiple choice	function words
Part 2		multiple choice	word formation	tenses/verb forms
Part 3		cloze test	cloze test	word formation
Part 4		function words		
Task	content words			

Table 1 Testing grammar and vocabulary in Croatia, Hungary, Slovakia and Slovenia

The discussion held in all these countries with secondary school teachers revealed that they had taught grammar and vocabulary a lot and it is an essential part of their syllabi. The school-leaving examination is considered a summative assessment and grammar and vocabulary are an inseparable part of teaching English. Sometimes teachers practice grammar more often than communicative language skills (as they believe that students might practice them outside the class) and want to get feedback on their teaching. It is generally thought that proficiency tests should be designed without language in use as these tests measure what the learner is able to do in the language in real-life situations.

	Croatia	Hungary	Slovakia	Slovenia
Task 1	-	9	20	13
Task 2	-	10	10	12
Task 3	-	9	10	10
Task 4	-	12	-	-
Task 5 in the reading section	8	-	-	-
Total	-	40	40	35

Table 2 The number of items in each language in use task in four countries

We can see that Hungary and Slovakia have the highest number of items related to assessing grammar and vocabulary. It might be a consequence of a mutual educational system for many centuries, on the other hand it might be influenced by the central European approach to education. Croatia was the only country, which does not have a separate language in use section, but included grammar in the reading section. On the contrary, it is necessary to emphasize that their testing vocabulary was based on testing content words, which can broaden the discussion if the assessors are completely prepared for other solutions made by their students.

4 TESTING ENGLISH IN SLOVAKIA IN 2014

To narrow the view on the school-leaving examination in English, it is necessary to look closer at the achievement of the students. The sample discussed in this paper is taken from the statistical analysis of English B2 in Slovakia. The psychometric characteristics of the test reveal that the number of the students was 14,198 achieving a mean of 61.5, standard deviation 17.7 and Cronbach's alpha - .919. The reliability of the test allows us to analyse it, focusing on single sections. To get the overall picture of the students' performances, the comparison of means for each section might help us.

	Listening	Language in use	Reading	Total
Mean	68.9	51.5	64.2	61.5

Table 3 Partial means versus total mean

Analyzing the achievements, it can be concluded that listening and reading measurements are quite close, the mean achieved in language in use is lower than the students' scores while being tested in receptive skills. On the other hand, the core distribution is highly peaked and symmetrical, since the distribution of scores above the central score is a mirror image of the distribution below the central core. On the other hand, score distributions referring to receptive skills are asymmetrical since the shapes of the upper and lower portions of the distributions are not mirror images of each other, and might be said to be skewed (Bachman 2005). These skewed distributions have their peaks toward one end of the distribution, having the longer tail at the lower end (negatively skewed distribution).

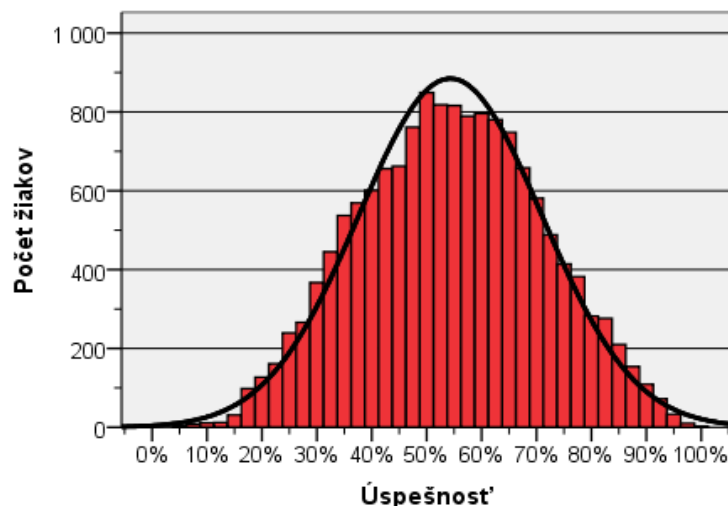


Figure 1 Distribution of scores achieved in the language in use section

Item analysis has revealed that most items of the multiple-choice task had been successfully achieved by students ranging from 44-95. The students had problems with word formation, mainly to derive the word inconvenient from the word convenience and tighten from tight, strategically from strategy. The cloze test included phrasal verbs (try out) and collocations (on the move), which made problems for most students.

The comparison of secondary-grammar school students' (SGS) with other school students' (others) performances might be significant as the former ones are obliged to sit a B2 test and it is elective of the latter ones as they can opt for B1. Testing scores achieved by students who could select an advanced level have revealed that good students' performances are not directly proportional to the type of school, but to their individual approach to language learning. There is only one item (24) in which secondary grammar school students achieved higher scores as the correct answer necessitated reading comprehension and the ability to recognize the time of the action. Underestimation of some items is usually based on improper reading or limited vocabulary of students. Some students seem to have problems with word collocations, phrasal verbs and tenses as they see those words or grammatical patterns in a new context. This problem with contextualized items might be explained by the fact that students link certain words and grammatical patterns to certain topics and are not flexible enough to guess the meaning of the word or improvise why using the particular word or structure.

Items	21	22	23	24	25	26	27	28	29	30
SGS	56.0	85.8	38.2	50.0	35.4	58.1	68.3	52.0	21.7	45.1
Others	73.6	93.6	44.1	48.8	44.1	63.9	80.6	56.5	26.4	53.2

Items	31	32	33	34	35	36	37	38	39	40
SGS	75.9	27.5	54.7	8.6	52.6	64.6	87.2	43.6	46.3	41.3
Others	86.0	29.1	56.2	16.4	65.6	77.3	95.0	54.2	40.8	58.5

Table 4 The difficulty level of multiple-choice items

The second task was based on testing word formation. In this case students did not achieve consistent scores and some items were more problematic for them than others. Analysis confirmed that some students had had problems to distinguish the proper part of speech; therefore the form of the word was not correct. On the other hand, the sentence structure is explained and trained properly while preparing students for writing tasks. As all students studying at the secondary grammar school are obliged to take English test B2, those who are not interested in languages achieve worse scores in word formation and cloze test.

Items	41	42	43	44	45	46	47	48	49	50
SGS	85.9	6.5	83.3	76.0	30.6	23.1	22.3	56.3	33.3	44.8
Others	89.0	8.7	85.6	79.3	31.4	40.8	31.8	68.6	46.2	54.2

Table 5 The difficulty level of word-formation items

The third task measured the usage of verb form and verb tenses in the contextualised text. The test-takers explained in their questionnaires that they had considered some items very difficult due to the fact that mostly they had practiced verb forms and verb tenses in a single isolated, discrete utterance. The students had not been trained in a gap-filling testing technique, which requires reading the text as a whole, allowing them to relate the information to extra-linguistic context and to interpret it accordingly. This provides a basis for the students to predict information for the gap, invoking the notion of expectancy. The type of information the students might be expected to supply could relate to linguistic form, semantic meaning and/or pragmatic use.

Items	51	52	53	54	55	56	57	58	59	60
SGS	74.0	29.8	84.1	28.1	81.1	62.4	33.8	69.4	63.4	47.7
Others	83.9	45.2	95.7	47.8	84.6	79.9	48.8	88.0	74.9	64.9

Table 6 The difficulty level of cloze-test items

The gap-filling task meets the philosophy of pragmatic expectancy grammar, which was defined by Oller (1979) as a psychologically real system that “causes the learner to process sequence of elements in a language that conform to the normal contextual constraints of that language, and ...requires the learner to relate sequences of linguistic elements via pragmatic mappings to the extralinguistic context” (Oller 1979, p. 38). Pragmatic expectancy grammar forces the students to integrate their knowledge of grammar, meaning and pragmatic use to complete the task. If the students are trained to control their grammar in isolated utterances without context, they are not prepared for controlling the grammatical patterns in a gap-filling task.

Bachman (1995) defined performance as the result of the test-taker's language knowledge interacting with the characteristics of the test task and other non-linguistic characteristics of the test-taker (i.e., his or her strategic competence, knowledge of the topic, affect and personal attributes). Purpura (2005) refers to grammatical performance as the observable manifestation of grammatical ability in language use. As McNamara (2000) notes, in the strong sense of performance assessment in language testing, the focus of this assessment is on the successful completion of a given task that requires language use, and not on the language use itself. The school-leaving examination tests related to writing and speaking use extended-production tasks that present input in the form of a prompt instead of an item. Extended-production tasks measure the student's ability to use grammatical forms to convey meanings in instances of language use and are scored with the rating-scale method. In Slovakia, analytic rating scales are used, which means that students' performances are rated on such features as task achievement (content), organisation, grammar and vocabulary. The four aspects are equally weighed and teachers consider analytic scoring very useful because it provides diagnostic information about the students' writing and/or speaking abilities. The analytic scales are used in the other three countries focused on in this article. The approach to the selection of criteria is influenced by the national educational system. Most countries use their analytic scales written in their mother tongues, as they are to be comprehensible, clearly stated, avoiding ambiguity. The reliability of the marking process is influenced by the budget the government invests into the objective way of assessment. Having not enough evidence about the objectivity, we came upon a decision not to compare the students' scores achieved in the language in use section with those achieved in grammar and vocabulary while the students were tested in writing and speaking. The data related to writing and speaking are rather confidential than public, and as they are recorded directly at schools, they are mostly unavailable.

5 CONCLUSION

The label CEFR has become so popular that world publishers label their material (course books, graded readers, grammar-reference books) without providing any evidence for their claims. Many decision-makers, education policy-makers put that label on national curricula only due to the fact that they used some descriptors in their documents. In one way, it can be considered as a promotion act, on the other hand, it can evoke disappointment as the document relates to learning, teaching and assessment and labelling the course books or curricula do not need to be labelled.

The CEFR (2001) is based on an action-oriented approach, so the descriptors for testing grammar are related to grammatical accuracy and those for testing vocabulary refer to vocabulary range and vocabulary control. On the other hand, summative evaluation (e.g. a school-leaving examination) provides test stakeholders with an overall assessment of a student performance related to grammatical ability and command of lexical repertoire as well as appropriate use of vocabulary. As communicative competence in the narrow sense includes linguistic, sociolinguistic and pragmatic competences, this more specifically language-related communicative competence is needed for the realization of

communicative intentions. The CEFR (2001) defines linguistic competence as knowledge of, and ability to use, the formal resources from which well-formed, meaningful messages may be assembled and formulated.

Whilst some tests measure grammar and vocabulary indirectly, it has been presented in this article that testing grammar and vocabulary meet the principles of the communicative approach. Teaching practice at schools influences the way grammar and vocabulary is assessed, vice versa communicative testing might influence the way grammar and vocabulary is taught. Testing grammar and vocabulary in the achievement tests using both perspectives: form-based as well as use-based enables one to obtain information on the students' knowledge of linguistic forms and usage in context, and a communication-based perspective in writing enables one to determine how well students can use linguistic forms to express a wide range of communicative meaning.

Although the CEFR (2001) adopted an action-oriented approach (as it was developed for adult learners to use target languages for practical reasons), testing grammar needs construct definition in which both grammatical form and grammatical meaning will be taken into account. Many applied linguists agreed that vocabulary is just one component of the assessment rather than the main focus. It is common practice to include testing vocabulary in the reading section of the test when the items are designed to measure students' ability to understand certain keywords or key phrases in context. In a controlled task of writing and speaking, students are judged on their ability to use vocabulary correctly and appropriately in relation to a particular task. In this case vocabulary use is specified as one of the assessment criteria and rated on a scale.

The inclusion of grammar and vocabulary components might be considered important as it can provide a lot of useful information regarding students' grammatical and lexical abilities. Whether or not to use a test to independently assess grammatical and lexical abilities depends on the purpose of the assessment. The choice of appropriate testing techniques is paramount due to the potential negative effect they might have on test performance. Any decision made in high-stake testing should be based on ethical awareness and fairness.

References

- Alderson, J C, Clapham, C & Wall, D 1996, *Language Test Construction and Evaluation*, 2nd edn, Cambridge University Press, Cambridge.
- Bachman, LF 1995, *Fundamental Considerations in Language Testing*, 3rd edn, Oxford University Press, Oxford.
- Bachman, LF 2005, *Statistical Analyses for Language Assessment*, reprinted, Cambridge University Press, Cambridge.
- *Common European Framework of Reference for Languages: Learning, teaching and assessment* 2001, Cambridge University Press, Cambridge.
- Oller, JW Jr 1979, *Language Tests at School: A Pragmatic Approach*, Longman, London.
- McNamara, T 2000, *Language Testing*, Oxford University Press, Oxford.
- Purpura, JE 2005, *Assessing Grammar*, Cambridge University Press, Cambridge.
- Read, J 2001, *Assessing Vocabulary*, Cambridge University Press, Cambridge.
- O'Sullivan, B 2012, 'A Brief History of Language Testing', in C Coombe, P Davidson, B O'Sullivan & S Stoyloff (eds.), *The Cambridge Guide to Second Language Assessment*, Cambridge University Press, Cambridge, pp. 9-19.