

## DEVELOPMENTS IN ESP TESTING PRACTICE: TEST DESIGN

Alla MĂMĂLIĞĂ

*Academy of Economic Studies of Moldova,  
Republic of Moldova, Chisinau, 61 Banulescu-Bodoni str.,  
Phone: +373 22 22 41 28, web site: [www.ase.md](http://www.ase.md)*

### Abstract

*Until the introduction of the Common European Framework of Reference for Languages, the assessment methods, and namely, the summative/achievement tests developed by the language teachers at the Academy of Economic Studies of Moldova and other universities, were not always designed according to certain common regulations/norms or were totally different from each other. Having the restrictions imposed by CEFRL, the content of the tests has improved significantly due to the application of new approaches not only to teaching, but also to language testing based on real-life tasks. Therefore, the tests for summative/achievement assessment, along with the teaching-learning methods have to be based on real-life situations, and close to current and, more important, students' future activities.*

**Key words:** *task-based language testing, traditional tests, assessment, test design, teaching methods.*

**JEL CLASSIFICATION:** A:23

### Introduction

This article aims to present some developments in testing English for Specific Purposes (ESP), that is Business English, in the Republic of Moldova within higher education institutions and namely Academy of Economic Studies of Moldova after the implementation of the Common European Framework of Reference for Languages: Learning, Teaching, Assessment (CEFR) in our country, using data which were collected by sharing the skills and knowledge of my colleagues and requirements imposed by the CEFR and the latest approach to teaching: task-based language teaching (TBLT) and, likewise, task-based language testing. The latter one has seemingly become a buzzword in our classrooms as it provides meaning to what should happen in a classroom and relates language use to some real-life settings. Similarly TBL testing is becoming a need and requirement imposed by the changes in language teaching methods and more demanding language learners.

As a result of all these our language instructors have become more aware of the different aspects involved in test design and administration, and have changed their ways of assessing students' performance. In what follows I will present the main findings of the study concerning the changes in ESP testing practice, which are supporting and complementing the general upgrading of ESP. I shall, therefore, focus on the data provided by the comparison of **task-based language testing** and **traditional language testing**, which illustrates most clearly the developments that have occurred in this area of ESP teaching due to the task-based language teaching approach.

With a view to offering a clear picture of the starting point, namely the kind of ESP tests which used to be administered to students in the early 2000s and before, I have outlined a traditional language test profile, based on a number of sample tests as well as on focus group discussions and interviews carried out in several higher education institutions and high schools within various workshops and seminars. The main characteristics I have identified are the following:

The traditional tests were mainly concerned with the testing of grammar and vocabulary either through specific test items or via writing essays on purposeless topics. The range of test methods used was rather limited, the most popular being translation both for grammar and vocabulary, defining terms, fill in the gaps, match and the like.

The testing of skills was neglected as a direct consequence of the absence of communicative methodology and materials from the English class. Particular interest could be noticed in the testing of specialist knowledge, rather than in the students' ability to use that knowledge when communicating in English in lifelike situations. Consequently, the texts and tasks were, with very

few exceptions, non- authentic and non-contextualised. In spite of the fact that essay writing and translation were widely used, the test authors showed no concern for developing marking criteria that might increase objectivity in these complex areas.

Test authors were not aware of the different kinds of tests, the only type being achievement, administered as end-of-year or/and end-of-term test. Usually, there was no common testing policy or co-operation in test design between teachers in the institutions considered.

The results of the study so far are that the developments in ESP testing have to be considered within the broader view of innovation and change in language teaching in the Republic of Moldova, where both training and new materials (i.e. textbooks) have had a strong impact on the teaching methodology and have affected testing substantially. The shift in focus on meaning and the communicative function of language used as a tool of TBLT, have turned the classroom into a springboard for real-life activities, enabling students to improve their language performance through realistic language tasks.

As mentioned above the TBLT outlines the importance of text and task authenticity, text and task relevance to learners' future activities. Thus, the following paragraph looks at the major changes that have occurred in testing ESP. ESP tests reflect this shift by using predominantly authentic texts from a wide range of sources such as articles in magazines and newspapers, business documents, interviews from the radio, lectures, etc. It is also important to note that the tasks give the students a real-life purpose by using the language in realistic situations and/or situations of professional relevance. This is achieved through contextualised tasks, which provide a context for using the language and a clear task environment, through more elaborate and explicit rubrics which describe the situation, the participants in communication, the roles, and the purpose. For example, *Leave a message on your English friend's answering machine that you would not be able to come to his/her party the following day.*

Thus, as far as the task type is concerned, the difference between the two kinds of tests, traditional and TBL tests is even more relevant, showing a dramatic change, in the sense that while traditional tests had practically no contextualized tasks and extremely few authentic ones, TBL tests have a large number of both authentic and contextualized tasks. To be noted, however, that non-authentic texts are still being used by some teachers.

As a consequence of the use of communicative/task-based tasks and of a change in teaching focus from grammar structures to skills, TBL tests show a move towards skill testing, with concern for all four skills. This represents the crucial difference from traditional tests, which mainly tested grammar and vocabulary (speaking, in particular, was very rarely, if ever, tested). This does not mean that TBL tests have abandoned grammar and vocabulary altogether, but they rather focus on all four skills, as well as on grammar and vocabulary.

*Grammar* is still assessed in TBL tests, but it is integrated into or combined with other skills such as reading, listening or speaking, compared with traditional tests, where each and every test included in the data set tested grammar. As a result of comparing more sample tests once used by me and my colleagues, I have found that, likewise, *vocabulary* testing occupies the second or the same, place in frequency in traditional tests, after grammar. With TBL tests, in my opinion, reading and writing come first, while grammar and vocabulary rank third and fourth, having a smaller importance. This development emphasises the trend towards the **testing of skills**. A significant change refers to the testing of *speaking* and *listening* which were absent altogether from the tests belonging to the traditional category. Giving the students the task *You are talking on the phone to a customer from France about a wrong delivery of goods. Some parts of the dialogues are missing. Compete it with the missing information.* assesses several skills at once: communicative based on TBL, writing, grammar, vocabulary and somewhere speaking, the crucial thing being the relevance of the task to real-life situations.

Another criterion used when analysing the two categories of tests, i.e. traditional and TBL-based, is the testing method. Thus, I have determined the fact that the range of testing methods is

much wider in the case of TBL-based tests. A variety of methods are used in such tests for testing different skills (*multiple matching, cloze, note taking, summarising, letter and essay writing, role play, e-mail/memo-writing* etc.). With traditional tests the testing methods are ‘overwhelmingly’ *translation, fill in the gaps with the correct form, sentence transformation*, which assess disembodied language elements, this being due to the fact that emphasis was laid on language usage rather than on language use, on accuracy rather than on fluency.

In order to illustrate this conclusion, let us take a closer look at the situation in the case of reading tests. The only testing methods used for traditional tests were multiple choice and comprehension questions, while for TBL tests there is a variety of testing methods, of which the most frequent are *true/false, matching elements, cloze, chart/table completion*.

The methods for assessing students’ **reading** skills show a focus on testing meaning comprehension, not memory. Methods like *paragraph reconstruction from jumbled sentences, structuring text into paragraphs, completing table/chart with relevant information* show that tests focus on processing and selecting information. *Matching* includes a variety of elements: *sentence to picture, heading to paragraph, multiple matching* but not *matching words with their definitions*. The presence of the *cloze* in reading tests indicates that integrated language is tested in TBL tests. The relatively big number of testing methods indicates the fact that TBL teachers are willing to innovate and try out a variety of methods.

As far as the testing of **writing** is concerned, I could say that this is the skill most frequently assessed by both categories of tests. Although the testing of this skill is widely present in the sample tests analysed, the range of methods used for testing writing is much wider in the case of TBL tests and there are considerable differences between traditional and TBL-based tests.

TBL-based tests involve writing **whole** texts with focus on discourse and register elements. These methods give a context and a purpose for writing and are relevant for professional purposes (*process description, project work, diagram interpretation in a full paragraph*), with less emphasis on specialist content and more on both communication and task completion. In the set of sample tests belonging to TOEFL, FCE, BEC or other international tests, the range of methods is fairly similar, with *essay and letter/memo/article/paragraph writing* in top positions, which means focus on testing writing of whole texts where organising information and message is the main marking criterion, as shown in the discussion of marking criteria below.

The traditional samples include tests which assess knowledge of specialist content (*definitions of specialist terminology, information transfer, answering questions on content*). There is also extensive testing of grammar through *translation, sentence completion, sentence formation from jumbled words, sentence completion, sentence writing based on a diagram* etc. A traditional test does not involve writing a professional document; whenever whole texts are produced it is through essay writing with no real purpose (e.g. “What are a manager’s functions”).

A great variety of methods is also used by TBL teachers in testing **listening**, a novelty in our tests at our university, as listening skills were not tested in hardly any of the old tests analysed. From now on testing listening becomes a requirement as it has been in all international language tests. The test methods used require students to concentrate on gist or specific information, rather than to understand every word. The most frequent method is *table completion*, which requires processing and selecting relevant information, but all other methods used such as *note-taking, matching pictures, matching elements*, focus on the message sometimes can integrate the testing of listening with that of grammar like in *reporting a conversation*.

As I have found from the sample traditional tests analysed, they do not offer examples of **speaking** tasks: nevertheless, the teachers who took part in group discussions on this topic over the seminars we have attended, mentioned that this skill used to be tested through monologue, by asking the students to speak on a given topic as it was customary at our university. The latest approaches to testing speaking, besides being a new development which is characteristic of TBL tests, also uses a great variety of methods. Although in our sample tests speaking is the least

frequently tested skill, the methodology of testing it is interesting and complex. The range of speaking skills tested is quite wide including *oral presentation skills on a real-life topic, debating, asking for and giving information, dialogue construction on a given topic, for and against discussion, etc.* Notable is student-student interaction (through *role play, conversation, debate* where the teacher assesses as an observer, without interrupting the students).

As far as the testing of **grammar** is concerned, TBL-based tests show a tendency to test it in corner, and in an integrative way, through the use of more elaborate techniques: *cloze, project work, completion of a dialogue, etc.* Unlike traditional tests, when: *translation, gap filling, multiple choice* are used mainly for testing grammar structures, TBL tests provide a **professional** context, which is both realistic and relevant for the learner as, for example, translating various office documents can be a task which the students will do in their future professional life. Thus, traditional tests focused on grammatical accuracy and metalanguage. Students were asked to explain/provide grammar rules, identify categories and/or to use them in non-authentic, non-contextualised tasks, having no real-life purpose, the rubrics being short and mechanical.

The tasks used to test **vocabulary** in TBL tests should be integrated with listening, reading, speaking, as simply writing such tasks as *word and phrase explanations, word families, word derivation* is similar to the ones used in traditional tests, and assessing vocabulary in isolation. It should be noted, though, that there is a strong element of vocabulary in context, focusing on language use *cloze, gap filling, labelling diagrams*, and identifying the meaning of words in context.

**Discourse elements** are tested implicitly in TBL-based tests. This can be seen from the marking criteria for writing (which include *coherence and cohesion, organisation and sequencing of ideas*), as well as from the methods testing other skills: listening (*note-taking, reporting conversation*), speaking (*oral presentation, for and against discussions, response to a given situation*) and reading (*text cohesion through sentence insertion in text, text reconstruction from jumbled paragraphs, joining sentences with cohesive elements etc.*). The tests focusing on discourse elements confirm the general trend towards testing integrative language. There was no such concern on the part of traditional tests, which tested discrete point items, especially through grammar and vocabulary.

## CONCLUSIONS

A TBL-based test profile comes to conclude the results of the analysis which have confirmed the fact that in-depth developments have taken place in the field of ESP testing of the Republic of Moldova since the implementation of the Common European Framework of Reference for Languages, which can be summarized as follows:

TBL-based tests reflect a concern for assessing the four skills through a variety of methods. Even though listening and, particularly, speaking are less frequently tested, they include a wide range of test methods and a variety of authentic texts from many different sources. These are new developments which were virtually non-existent in traditional tests.

TBL-based tests show a move towards assessing the students through authentic and contextualised tasks. Students are not only required to make full use of their language competence, they are given a real-life purpose as well, as the test items focus on meaning and the communicative function of language with a task-fulfilment outcome.

However, old testing strategies and methods can be still identified in TBL tests in varying degrees. The analysis of sample tests language instructor uses has revealed that, along with the four skills, vocabulary and grammar are still tested, due to requests from students, and to the fact that grammar and vocabulary tests provide a clear track record of progress for both teachers and learners.

## **BIBLIOGRAPHY**

1. BURGER, N. C., *Focus on form in task-based instruction: an exploratory study*. København. 2001. ISBN: 9789462599215
2. NUNAN D., *Task-based language teaching*. Cambridge University Press, 2004, ISBN-13: 978-0521549479
3. ROD E., *Task-based language learning and teaching*. Oxford: Oxford University Press, 2003, ISBN: 0194421597
4. LONG, M. H., & DOROUGHTY, C., *The Handbook of language teaching*. Malden, MA: Wiley-Blackwell, 2009, ISBN: 978-1-4051-5489-5.
5. WILLIS, J., *A framework for task-based learning*. Harlow: Longman. (1996). ISBN-10: 0582259738;
6. WILLS, J. *A Framework for Task-Based Learning*. Longman OUP 2007