

# THE LIMITS OF ESP TESTS

Norica-Felicia BUCUR\*  
Camelia NEAGU\*\*

## Abstract

*Global market forces have determined not only higher education institutions all over the world to include ESP courses in their curriculum to enhance their students' future employability, but also public and private organisations to offer their employees the opportunity to attend ESP courses in order to meet the continuously growing ESP needs. From this perspective, ESP competence could become a subcomponent of one of the key competences for lifelong learning, communication in foreign languages. Therefore assessing ESP competence seems to acquire paramount importance since stakeholders need accurate information about the ESP learners' abilities to cope with specific language tasks. This article offers a concise overview of the principles and practices of ESP assessment, a detailed description of the features of ESP tests, while focusing particularly on the limits of ESP tests in order to identify possible solutions to overcome them.*

**Keywords:** *English for Specific Purposes, language assessment, ESP assessment, language test, ESP test, specific background knowledge, language proficiency, authenticity, specificity, tasks.*

## 1. Introduction

Nowadays, no formal learning can be envisaged without some form of assessment, because educational stakeholders are eager to be informed about students' progress. Moreover, the course objectives, content and methodology are also scrutinized in this process, so that suggestions for improvement could be made, if necessary.

English for Specific Purposes (ESP) is specialized English language teaching that aims to develop the specific skills of the learner in response to the needs identified or indicated by various stakeholders. This specificity includes equipping learners with 'not only knowledge of a specific part of the English knowledge, but also competency in the skills required to use this language' (Orr, 2002: 1). Furthermore, ESP learners comprise almost all adult age groups, as well as cultural, linguistic, professional and academic backgrounds.

In line with late 20<sup>th</sup> century trends characterising curriculum development in general and English as a Foreign Language curriculum development in particular, ESP practice has evolved into a spiraling protocol of standard procedures: needs analysis; syllabus design; selection and production of materials; methodology; assessment. Thus, ESP practitioners play multiple roles, being responsible for designing customized syllabi, preparing materials, carrying out research and evaluating specific language progress.

The field of ESP assessment has been seen as a separate and distinctive part of a more general movement of English language assessment, focusing on measuring specific uses of English language,

among identified groups of people. Moreover, ESP assessment has been viewed in the broader context of the teaching and learning process. Thus, assessment does not stand alone, but occupies a prominent place in the ESP process, giving an ESP teacher a wealth of information on the effectiveness and quality of learning and teaching (Dudley-Evans and St. John, 1998: 121). On the other hand, tests enhance the learning process and act as a learning device, and, particularly, an ESP test is an aid to learning, encompassing benefits such as reinforcement, confidence building, involvement and building on strengths (Dudley-Evans and St. John, 1998: 210-212).

This article reviews the theoretical issues related to ESP assessment in general, trying to pinpoint to the characteristics of ESP tests. The limits of ESP tests, as evinced from research conducted in this field, are briefly described and possible solutions are indicated.

## 2. ESP Assessment: A Brief Outline

The process/product dichotomy helps us differentiate between the concepts of assessment and testing, in general: assessment is a process, which may offer answers not only to questions such as how much students have learned (this purpose is achieved by means of a test = product), but also to questions on how they learned and why certain results occurred. This information is obtained through techniques other than tests, including observations, surveys, interviews, performance tasks, and portfolios. Thus, assessment is a comprehensive concept, centering its endeavors on student learning, and serving the purpose of student improvement and development through a variety of ways.

---

\* Assistant Lecturer, PhD, Faculty of Social and Administrative Sciences, "Nicolae Titulescu" University of Bucharest (e-mail: feliciab@univnt.ro).

\*\* Lecturer, PhD, Faculty of Social and Administrative Sciences, "Nicolae Titulescu" University of Bucharest (e-mail: camelia.neagu@univnt.ro).

The guiding principle of assessment in ESP pedagogy is gathering evidence to understand the effectiveness of the course in terms of the skill enhancement of learners. The traditional needs analysis in ESP covers the purpose of the assessment, the personal, educational, and knowledge characteristics of the learners, and the context of specific purpose language use (Douglas, 2013: 367).

Therefore, assessment instruments employed in specific purpose programs are not, in fact, completely different from assessment instruments in general or from assessment tools used in language programs as such, because the ultimate purpose of assessment is to give learners an opportunity to show what they have learned and what they can do with what they have learned, by being given the same instructions and the same input under the same conditions.

Nevertheless, ESP assessment is distinguishable from assessment in general and language assessment in particular as it targets specific purpose language abilities. Thus, if we want to know how well individuals can use language in a specific context of use, the assessment needs to include both (1) their language knowledge and their background knowledge and (2) their use of strategic competence in relating the salient characteristics of the target language use situation to their specific purpose language abilities (Douglas, 2000: 282).

### 3. Main Characteristics of ESP Tests

From a more general perspective, tests are summative assessment tools and their goal is to evaluate student learning by comparing it against some standard or benchmark. In contrast with formative assessment instruments, which exhibit high degrees of flexibility, 'most tests are conducted under supervision and require candidates to answer questions in a given time limit, without reference to books or other people' (Dudley-Evans and St John, 1998: 211). What is more, tests can be labeled as low-stakes or high-stakes, depending on their function. In a low-stakes test, the results typically matter far more to an individual teacher or student than to anyone else and it is usually distributed, rated and graded by the teacher personally, whereas in a high-stakes test, test results are used to determine an important outcome, having major consequences or being the basis of a major decision.

Language tests play a powerful role in many people's lives, acting as gateways at important transitional moments in education, in employment, and in moving from one country to another (McNamara, 2000: 4). As compared to general English tests, ESP tests are developed within an ESP context, aiming to comply with the narrowly defined requirements of any specific area of language use.

What really matters in ESP testing is whether learners can communicate in a specific target language and use knowledge of the field in order to achieve their aims, in order to understand and be understood, in

order to get their message across in English. The ESP approach in testing is based on the analysis of learners' target language use situations and specialist knowledge of using English for real communication. Thus, when envisaging an ESP test, the ESP underlying principles should pervade the whole process: (1) language use varies with context; (2) specific purpose language is precise; (3) there is an interaction between specific purpose language and specific purpose background knowledge (Douglas, 2013: 368).

ESP tests are not exclusively employed to measure proficiency, but they may also be given to place students or to check their progress. These functions may even overlap in certain situations (Hutchinson and Waters, 1987: 146), offering valuable information to all stakeholders. Therefore, a placement test could not only reveal the learner's needs, but also his/her potential for learning along the ESP course, and an achievement test could not only give the learner a glimpse on his/her progress, but it might also function as a motivator.

Moreover, ESP tests are criterion-referenced tests, being designed to represent the candidate's level of ability, hence his/her performance is interpreted with reference to the criterion judged to be essential for proficiency in a particular task (Hutchinson and Waters, 1987; Douglas, 2000). Thus, an ESP test calls for an interaction between the test taker's language ability and specific purpose content knowledge, on the one hand, and the test tasks on the other (Douglas, 2000: 19).

Another important aspect of ESP testing is the target group. ESP learners stand out as it is assumed that they have become aware of both their target needs (subdivided into necessities, lacks and wants) and their learning needs (Hutchinson and Waters, 1987: 54-61). Unlike general English tests, which are sometimes targeted for certain age groups (e.g. Cambridge YLE – young learners), ESP tests are more likely to be used with adults, either at a tertiary level institution or in a professional work situation.

Although the washback effect is not ESP specific, it is worth discussing this issue here, since it heavily weighs on test quality. At macro level, the higher the stakes, the more impact a test may have 'on learners and teachers, on educational systems in general, and on society at large' (Hughes, 2003: 53). At micro level, washback refers to the impact of testing on 'what is taught and how it is taught' (Dudley-Evans and St John, 1998: 214). At this level, washback may prove beneficial in certain situations, as it could show the teachers either what should be changed in the tests and, the opposite, if some parts of their teaching should be changed in order to focus more on what it is tested. Moreover, the changes operated in the tests by the teachers could help learners improve their language abilities.

When devising an ESP test, valid and consistent measures of language ability should be developed as

well. These measures are expected to be as reliable and authentic as possible, provide accurate descriptions about specific language abilities, have beneficial impacts, be practical and cost effective in terms of administration, time, money and personnel. Nevertheless, only authenticity, interactiveness and impact/practicality are quintessential for language tests in general (Bachman and Palmer, 1996), and the quality of any ESP test stands out especially to authenticity (Dudley-Evans and St. John, 1998, Douglas, 2000).

Considering the most important factor affecting the quality of an ESP test – the level of *authenticity*, ideally, an ESP test should engage the test takers in accomplishing various genuine tasks through which their general English knowledge (linguistic competence) could interact with their ESP content knowledge in a real life context. Such an interaction is deemed absolutely necessary throughout the entire ESP testing process, as authenticity of the tasks refers to the similarity of the task content to the specific and specialized target language situation. In addition to that, authenticity is important because of its potential effect on the test takers' perceptions of the test and, hence, their performance (Bachman and Palmer, 1996: 24).

In a perfect world, after attending an ESP course, learners' previous English skills would refine, helping them to successfully deal with domain specific situations. Hopefully, their background knowledge would increase and this would enable them to make appropriate inferences and thus boost their ESP skills. Tests taken by learners at the end of their ESP course should provide stakeholders with some evidence, by predicting how effective and confident the test-takers might be in a specific target situation.

#### 4. The Difficulties of ESP Testing – Possible Solutions

Even if ESP is widely acknowledged as a well established field (see Hutchinson and Waters, 1987; Dudley-Evans and St John, 1998; Douglas, 2000; Basturkman, 2010), when it comes to understanding what is to be assessed in an ESP test, practitioners seem to be on less solid ground. Two theoretical issues need considering in Douglas's view (2000, 2013): (1) whether the construct of specific purpose language ability actually exists and (2) how the criteria for assessing specific purpose language performances should be derived.

On the one hand, it is argued that the only justification for developing specific purpose language tests is to achieve face validity, and test developers usually achieve this by exclusively focusing on the purely linguistic elements of specific purpose communication (Davies, 2001: 143). On the other hand, stemming from the absolute and variable characteristics of ESP (Stevens 1988 in Dudley-Evans and St John, 1998: 3), it is argued that a specific

purpose language performance depends on non-linguistic knowledge (Jacoby and McNamara, 1999: 233). In reality, these two opposing views are brought together on a continuum, because ESP tests cannot acquire absolute values in terms of authenticity and specificity. Thus, ESP tests range from low authenticity/specificity to high authenticity/specificity, depending on specific context itself (Douglas, 2013: 371).

Authenticity does not lie in the mere simulation of real-life texts or tasks, but rather in the interaction between the characteristics of such texts and tasks and the language ability and content knowledge of the test takers (Douglas, 2000: 22). Therefore, the distinction between *the ability to do future tasks or jobs in the target language use situation* and *the ability to use language in specific future tasks or jobs* is crucial, as any relevant ESP test should find a way to successfully combine the assessment of these abilities.

The problem related to properly defining specific purpose language ability proves to be so complex that it is really difficult to come up with a solution that would comply with the demands of any ESP situation. Nevertheless, if further analyzing this issue helps identifying the elements that define specific purpose language ability: *context* and *tasks*. According to Douglas (2000: 41), in producing reliable ESP tests, 'we need an understanding, first, of the context and tasks in the specific field we are interested in, and second, of how to translate these features into test tasks'. The framework proposed by Douglas (2000) for analyzing target language use and test task characteristics is meant to provide a basis for test development, as well as for analyzing existing specific purpose language tests to determine their degree of specificity. The framework is made up (1) *rubric* – 'characteristics that specify how test takers are expected to proceed in taking the test' (Bachman, 1990, in Douglas, 2000: 50); (2) *input* – 'the specific purpose material in the target language use situation that language users process and respond to' (Douglas, 2000: 55); (3) *expected response* – 'refers to what the test developers intend that the test takers do in response to the ESP situation they have attempted to set up by means of the rubric and input' (Douglas, 2000: 62); (4) *interaction between input and response* – in terms of reactivity, scope and directness (Douglas, 2000: 63); (5) *assessment* – 'consists of a construct definition, a set of assessment criteria, to be called criteria for correctness, and a set of procedures for rating or scoring the performance' (Douglas, 2000: 67).

The degree to which the learners' level of language knowledge and their level of specific background knowledge affect their ESP test performance has been intensively investigated (Salmani Nodoushan, 2007; Krekeler, 2006; Clapham, 2000, 1996; Ridgway, 1997). Nevertheless, research findings have been unable to tip the balance in favour of any of these variables, as, in some cases a specific background knowledge effect is detectable and in

others it is not. For example, even if intermediate proficiency test takers seemed to have benefited from their background knowledge to obtain a higher score, for low and high proficiency candidates the value of this variable was irrelevant (Clapham, 1996, 2000).

This lack of conclusiveness related to the effect of specific background knowledge on ESP test performance has been attributed to four problems: (1) background knowledge is difficult to assess; (2) there is lack of criteria to identify qualities of specific language material, which makes it difficult to compare relevant studies; (3) there may be considerable task effects and the test tasks also vary to a great extent across studies; (4) there seem to be interaction effects between background knowledge and language proficiency (Thoma, 2011: 115).

## 5. Conclusions

The field of ESP language testing has a rich and dynamic future. The troublesome issues in ESP

language testing have been briefly commented upon in this paper and partial solutions have been indicated. The problem of authenticity/ specificity, which could apparently be solved by selecting input data from genuine sources, is difficult to cope with in reality, as low and high levels of authenticity/ specificity could exist for ESP tests, depending on the situation. Moreover, so far research findings have not provided irrefutable evidence related to the importance of specific background knowledge when taking an ESP test (Thoma, 2011: 118).

Therefore, more research is badly needed to indicate practical solutions to the problems evinced by specialists investigating the field of ESP testing. Experts in ESP teaching should involve professionals from various domains in designing and conducting relevant research, so that ESP tests would become more fair, reliable, authentic, valid and practical.

## References

- Bachman, L.F.; Palmer, A.S. (1996). *Language Testing in Practice: Designing and Developing Useful Language Test*. Oxford: Oxford University Press
- Basturkman, H. (2010). *Developing courses in English for specific purposes*. London: Palgrave Macmillan.
- Clapham, C. (1996). *The development of IELTS: a study of the effect of background knowledge on reading comprehension*. Cambridge: Cambridge University Press.
- Clapham, C. (2000). Assessment for academic purposes: Where next? *System*, 28, 511–521.
- Davies, A. (2001). The logic of testing languages for specific purposes. *Language Testing*, 18: 133 – 148.
- Douglas, D. (2000). *Assessing Languages for Specific Purposes*. Cambridge: Cambridge University Press.
- Douglas, D. (2013). ESP and Assessment. In Brian Paltridge and Sue Starfield (eds.) *The Handbook of English for Specific Purposes*, 367-383, Oxford: Wiley-Blackwell
- Dudley-Evans, T.; St John, M. (1998). *Developments in ESP. A multi-disciplinary approach*. Cambridge: Cambridge University Press.
- Hughes, A. (2003). *Testing for Language Teachers. Second Edition*. Cambridge: Cambridge University Press.
- Hutchinson, T.; Waters, A. (1987). *English for Specific Purposes. A learning-centered approach*. Cambridge: Cambridge University Press.
- Jacoby, S.; McNamara, T. (1999). Locating Competence. *English for Specific Purposes*, 1: 213 – 241.
- Krekeler, C. (2006). Language for special academic purposes (LSAP) testing: the effect of background knowledge revisited. *Language Testing*, 23: 99-130.
- McNamara, T. (2000). *Language Testing*. Oxford: Oxford University Press.
- Orr, T. (ed.) (2002). *English for Specific Purposes*. Alexandria: TESOL
- Ridgway, T. (1997). Thresholds of the Background Knowledge Effect in Foreign Language Reading. *Reading in a Foreign Language*, 11(1): 151-168
- Salmani-Nodoushar, M.A. (2007). Are Task Type And Text Familiarity Predicators of Performance on Tests of Language For Specific Purposes? *The Asian ESP Journal*, 3: 67-96.
- Thoma, D. (2011). Strategic Attention in Language Testing. *Metacognition in a Yes/No Business English Vocabulary Test*. Frankfurt am Main: Peter Lang