Initial testing in English for upper secondary students in Romania

Letitia Trif a *, Teodora Popescu b

a 1 Decembrie 1918 University of Alba Iulia, 5, Gabriel Bethlen Str., Alba Iulia 510009, Romania
b 1 Decembrie 1918 University of Alba Iulia, 5, Gabriel Bethlen Str., Alba Iulia 510009, Romania

Abstract

The aims of this study are to assess the performance of upper secondary school students studying English as a foreign language (FL1 and FL2) following the initial testing, compulsory for all secondary school students starting with school year 2011-2012, and to establish remedial action to be taken in order to address the mixed ability situations and the large gap that often exists between the level imposed by national curricula and the actual students’ level. The initial tests design was checked against the national curriculum standards. We randomly analysed the initial tests taken by 51 EFL1 students (10th grade) and 49 EFL2 students (10th grade). Testing items were clustered around the following competences and sub-competences: I. Reading comprehension (reading for detail, opinion/attitude); II. English in Use (structure, lexis); III. Writing (letter writing). Pedagogical insights into classroom teaching methodology to address the existing problems facing state schools in Romania will be provided.

1. Introduction

English language education in Romania is an extremely diverse landscape, given the disparity between levels of ability within one class or school as well as the variety of teaching methods used by different teachers. Despite a national curriculum aligned with the Common European Framework of Reference for Languages, the actual classroom situation is very different from the learning objectives and outcomes provided for by the National Council for Curriculum in Romania. Curricular reform in Romania started in 1998, with the elaboration of The National Curriculum. Framework Plan for Primary and Secondary Education, by setting a national curriculum, syllabi and approved teaching manuals, for all school subjects.

Alternative manuals started to be used by different school, teachers and headmasters being the decision makers in this sense. In general terms, English manuals introduced starting with 1998, were published in England, or in Romania, but under the supervision of British manual writers. Compared with traditional manuals, which were based on the grammar-translation method, the new ones were designed in a more communicative manner, focusing on all four language skills: reading, writing, speaking and listening. Likewise, the teaching methods also had to change: from a more-teacher-oriented approach to a student-centred one. The activity of the British Council at that

* Corresponding Author Trif Letitia (*). Tel.: + 40-0258-813722
E-mail address: trifletitia@yahoo.com
time was extremely beneficial in terms of teacher training – many English teachers benefited from refresher courses, and many had the opportunity to attend ELT methodology training courses. The introduction of the communicative method unfortunately in some cases superseded the teaching of grammar altogether, and in many cases teachers ended up with students who were able to express their ideas in speaking, but could not do this in writing. Other teachers remained faithful to their traditional methods, and continued to teach grammar in an intensive and explicit manner.

2. Literature Review

Our endeavour needs to start from a review of classroom testing and assessment, as we wanted to focus on the first nation-wide initial English testing ever used in Romania for upper secondary students. By ‘validity’ in testing and assessment we have generally referred to finding out whether a test ‘measures accurately what it is intended to measure’ (Hughes, 1989, p. 22), or discovering the ‘appropriateness of a given test or any of its component parts as a measure of what it is purposed to measure’ (Henning, 1987, p. 170). Understanding validity in this way means that when we design a test we have an intention to measure something, which ‘something’ is ‘real’, and that research into validity aims at finding out whether a test ‘actually does measure’ what is intended.

Traditionally, in specialist research, validity was divided into three different ‘types’, each being related to the kind of evidence that would count towards proving that a test was valid. According to Cronbach and Meehl (1955) these types were: 1) Criterion-oriented validity (Predictive validity and Concurrent validity); 2) Content validity; and 3) Construct validity.

Criterion-oriented validity refers to the situation in which the tester is interested in the relationship between a particular test and a criterion to which we want to make predictions. Validity evidence is therefore the strong point of the predictive relationship between the test score and the performance on the criterion.

Content validity aims at proving whether the content of the test is a representative sample from the area that is to be tested. Therefore, it is usually done using expert evaluators, who may be subject teachers, or language teachers who have many years’ experience in teaching English for special purposes. Evaluators look at texts that have been selected for inclusion on the test and check them for their representativeness within the content area. Next, the items used on the test need to lead to responses to the text, from which we can make assess the test takers’ ability to process the texts.

Construct validity represents the extent to which operationalisations of a construct (i.e., operational measurable and observable, interrelated concepts) actually measure what the theory states they do.

Validity theory has emerged from the practice of large-scale testing and assessment, i.e. situations in which a test provider develops a test that is used on a national or even international scale. In this case the test provider needs to ensure that the use of the test is appropriate for its declared purpose and target population. This refers to investigating the extent to which score meaning can be generalized beyond the test conditions, and to the ability to communicate in non-test situations. What needs to be considered, however, is the fact that teachers of English are not really trained into the difference between classroom assessment and large-scale assessment. Textbooks, which may display a much more pedagogic orientation, feature nevertheless a tendency to regard classroom assessment in terms of traditional categories such as ‘placement’ (relating to establishing needs), achievement of proficiency as part of programme evaluation, and progress tests to inform lesson development (Bailey, 1998, p. 39). This results in a presentation and knowledge of norm-referenced statistics, which not many school teachers are conversant with. Therefore, regardless of the usefulness of the practical guidance in texts for teachers, the difference between the classroom and large-scale testing is not taken into account, creating much confusion among teachers.

3. Research Methodology

The design of the initial tests was checked against the national curriculum standards. We randomly analysed the initial tests taken by 51 students of EFL1 (10th grade) and 49 students of EFL2 (10th grade). Testing items were clustered around the following competences and sub-competences: I. Reading comprehension (reading for detail,
opinion/attitude); II. English in Use (structure, lexis); III. Writing (letter writing), therefore the initial tests taken were written, and assessed reading and writing skills, as well as grammar and vocabulary. There was no test to assess listening and speaking skills. As the test was taken nation-wide at the beginning of the 10th grade, we resorted to an analysis of the syllabus for the previous school year, which the students should have covered fully. The national curriculum for the 9th grade is structured into the following parts: A. General competences (1. Reception - reading and writing in different communicative situations; 2. Production – Oral and Written; 3. Interaction – Oral and Written; 4. Transfer and mediation of oral and written messages). These general competences are then detailed into B. Specific competences and C. Presentation of content (for both core and differentiated curriculum).

The test for EFL1 consisted of 8 tasks, grouped as follows: Reading Comprehension (I, II and III), English-in-use – Vocabulary (IV and V); Grammar (VI – multiple choice; VII – rephrasing) and Writing (VIII). Here is the writing task, an informal letter to a friend, based on narration:

VIII. Composition. Imagine that you were Robert. Write a letter of about 100 words to a friend, explaining the trick which the old man played on the crowd. Use your own words as far as possible.

30 points.

Figure 1. Exercise VIII of Initial test EFL1, 10th grade

The test for EFL2 consisted of 2 parts, the first one testing Reading comprehension and English-in-use, while the 2nd part tested the Writing production ability of students. For the sake of exemplification, here is the reading comprehension text, consisting of 5 multiple choice items.

PARTEA I

1. Read the text below. Are the sentences 1-5 “Right” (A) or “Wrong” (B)? If there is not enough information to answer “Right” (A) or “Wrong” (B), choose “Doesn’t say” (C).

When I started running seven years ago, I could manage only about a quarter of a mile before I had to stop. Breathless and aching, I walked the next quarter of a mile, then I jogged the next quarter of a mile, alternating these two activities for a couple of miles. Within a few weeks I could jog half way round Hampstead Heath without stopping. Soon I started to run up the quarter-mile slope to the top of Parliament Hill, although I had to stop at the top to get my breath back. Eventually I found that I could even manage to get up the hill comfortably. I started to run because I felt desperately unfit. But the biggest pay-off for me was-and still is- the deep relaxation that I achieve by taking exercise. It tires me out but I find that it does calm me down.

1. The text is about running.
   A. Right  B. Wrong  C. Doesn’t say
2. After one week he could jog half way round Hampstead Heath without stopping.
   A. Right  B. Wrong  C. Doesn’t say
3. He started to run because he felt unhealthy.
   A. Right  B. Wrong  C. Doesn’t say
4. He didn’t run alone.
   A. Right  B. Wrong  C. Doesn’t say
5. He achieves deep relaxation by taking exercise.
   A. Right  B. Wrong  C. Doesn’t say

(20 points)

Figure 2. Part I of Initial test EFL2, 10th grade

The results obtained by students were analysed quantitatively and qualitatively, and we tried to identify possible reasons why students obtained these scores. We further on tried to adopt a global perspective and suggest pedagogic approaches that may improve the situation in the cases that call for remedial action.

4. Results and Interpretation

To start with, the testing process itself did not concentrate on all the learning objectives and outcomes of the national curriculum. The speaking production and the oral and written interaction were not tested at all. In terms of content presentation, the test items for the reading comprehension are well constructed. They display an intermediate difficulty, appropriate for the level that students needed to have achieved by the end of the 9th grade and beginning of the 10th grade, in both cases, EFL1 (7 years of English study) and EFL2 (5 years of English study). Mention should be made of the fact that the tests that we focused on and analysed in our research were administered at a technological school. Tests were not differentiated according to the upper secondary school profile.
The test administered to EFL1 students, based on our sample, revealed the following information: a global correctness percentage of 41.10%, which is below the normal pass level according to the Romanian grading system (5 is the pass grade). In the case of Reading Comprehension, the percentage was 28.30%, extremely poor in terms of students’ mastering the reading comprehension learning strategies. English-in-Use recorded the highest percentage: 45.91%, while Vocabulary hit 32.5%. The lowest score was in the case of Writing, i.e. 22.88%. The average recorded below the pass score for the Romanian educational system is indeed worrying. It proves that students have insufficiently acquired the communicative competence so far and raises serious questions as to whether the current manuals in use would be of any help any more. It also reveals that the curricular objectives set at national level are too high for the actual class reality.

In the case of the test administered to EFL2 students, the global results proved insufficient language competence, featuring a correctness percentage of 41.73%. On the Reading Comprehension part, students scored the best, which proves that they have had sufficient practice in this type of learning strategies. 13 students out of 49 scored the maximum. The situation was however different in the case of the multiple choice English-in-use exercise, in which 1 student scored 0 points, and none scored 20 points (the maximum). It was even worse in the case of the error-correction English-in-use exercise: 31 students scored 0 points and no student scored either 18 or 20 points. This reveals a poor knowledge of the students in identifying correct/incorrect grammar structures. In the case of writing production, the results displayed a clear lack in previous training for the writing skill. Worryingly enough, there were cases in which many students (32 out of 49) did not write a single word for the writing task. The maximum score obtained by students on this task was 22 points out of 30 (2 students only).

The poor results obtained by students may also be indicative of the fact that in this case, English being studied as a second foreign language, students do not have enough interest in developing their communicative competence (at least in terms of written reception and written production). Motivation (both intrinsic and extrinsic) may also be low.

---

Figure 3. Part I, ex. I, Initial test EFL2, 10th grade

Figure 4. Part I, ex. II, Initial test EFL2, 10th grade

Figure 5. Part I, ex. III, Initial test EFL2, 10th grade

Figure 6. Part II, Initial test EFL2, 10th grade
5. Conclusions and Recommendations

Initial testing results were centralized and analysed at the level of each School Inspectorate. The observations were as follows:

The lowest scores were recorded in lower secondary schools and technological upper secondary schools. Most of the students do not possess basic concepts concerning English grammar; display a low written expression capacity; do not know tenses, make many spelling mistakes and do not use tenses correctly (sic!). In order to take remedial action, teachers of English will do the following:

- adapt learning content to each student’s understanding peculiarities;
- make handouts and establish individualized learning tasks;
- revise grammar elements previously studied;
- apply interactive, student-centred methods;
- initiate and hold more conversation classes and assign writing tasks on given topics.

For students in the last upper secondary school grades, who will take the final language proficiency test before the school leaving examination, teachers will organize a refresher programme to revise basic concepts and notions, through individual work and through weekly classes organised in schools.

One thing worth mentioning in this context is exactly what we mentioned at the beginning of this paper, i.e., teachers often do not make the difference between classroom and large-scale testing. Indeed, initial testing was extremely useful as a diagnostic tool, as it revealed the actual state of language learning in Romanian schools. However, as we said before, the tests that we focused on were taken by students from a technological upper secondary school. There is no difference in the national curriculum for languages for the different types of upper secondary education (theoretical, technological, vocational/arts and crafts). From a macro-educational perspective, this generalization is counterproductive. As long as there are different types of upper secondary schooling, the curriculum should be likewise differentiated. There is little chance that the communicative competence of students whose average was under the national pass grade will dramatically improve their level of language proficiency in the coming years. Setting realistic and achievable learning objectives and outcomes would probably be the better option. Nothing has been mentioned in terms of teaching manuals. This is one important element that should not be overlooked, as the choice of teaching/learning materials is of crucial importance in stimulating students’ motivation and involvement. Therefore, English for Specific Purposes textbooks should also be accepted in technological or vocational upper secondary schools.

References