

# The Impact of the Common European Framework of Reference on Teaching and Testing in Central and Eastern European Context

Jana Bérešová  
Trnava University - Slovakia  
jberes@truni.sk



Synergies Europe n° 6 - 2011 pp. 177-190

Dates de soumission/acceptation : 3 mai - 4 septembre 2011

**Résumé :** L'article étudie l'impact du Cadre européen commun de référence (CECR) sur l'enseignement et l'évaluation de l'anglais dans les pays de l'Europe centrale et orientale (connus dans le domaine de l'économie comme les Pays avec une économie de transition, ou dans le domaine de l'histoire, les anciens pays socialistes). Ceci est fait en se reportant aux changements brusques dans le domaine de l'enseignement des langues, ce qui a pour conséquence la nécessité d'une réforme de l'examen de fin d'études dans la plupart d'entre eux. L'introduction du CECR a suscité un vif intérêt parmi des professionnels de l'enseignement des langues en Europe centrale et orientale afin d'établir une correspondance entre les programmes d'études et les syllabus avec les niveaux de référence pour établir les objectifs. La gamme de compétences différentes décrite dans le CECR a servi de ressource aux enseignants des langues pour élargir la gamme de leur enseignement. L'article examine la réforme de l'examen de fin d'études en Slovaquie en se référant à une part d'un programme de recherche, mené en 2008, étudiant les premiers essais de faire correspondre l'examen national d'anglais au CECR.

**Mots-clés :** Cadre européen commun de référence, Manuel, l'impact sur le programme d'études national et la réforme de l'examen de fin d'études national

**Abstract:** The article explores the impact of the Common European Framework of Reference (CEFR) on teaching and testing English in countries of Central and Eastern Europe (referred to economically as Countries with Economies in Transition, or historically, the former socialist countries). It does so by reporting on the abrupt changes in language teaching, resulting in the necessity of school-leaving examination reform in most of them. The introduction of the CEFR fostered a genuine interest of language professionals in Central and Eastern Europe to relate the curricula and syllabi to the common referential levels for setting objectives. The range of different competences described in the CEFR provided a useful reminder to language teachers to extend the range of their teaching. The article considers the school-leaving examination reform in Slovakia by reporting on part of a research project, conducted in 2008, on the first attempts of relating the national examination in English to the CEFR.

**Keywords:** Common European Framework of Reference, Manual, impact on the national curriculum and school-leaving examination reform

## 1. Introduction

Before 1989, language teachers in Central and Eastern Europe were influenced by a traditional approach to teaching. The teacher was the authority responsible for selecting the correct answers, conducting, guiding and controlling the students' behaviour in the target language. The goal of the lesson was mostly a linguistic item, explained by the teacher and compared to the students' mother tongue. Students were expected to memorise the word lists by heart without any context. Reading and listening were skills trained without practical objectives; students were mostly expected to retell the content of reading or listening material. The interaction was mostly teacher-directed, questions were posed by the teacher and students were expected to answer what they had memorised. Students could generally learn only what their teachers knew, and knowledge was more important than the real usage of the target language. This system of teaching and assessing was strictly controlled through the curricula, a single textbook and regular visits to the school by the inspectors. These inspectors were typically not even graduates in the target languages, and therefore they were more interested in checking the methods and various educational aspects than the quality of target language used in class. Many language teachers with degrees in the language they taught had never been exposed to real usage of the language they studied. They taught what they had learned.

The old system of assessment - testing knowledge more than language communicative competence - was not external, valid or reliable. Writing and speaking were tested. The topics for writing were set centrally, but usually marked internally in schools by the students' own teachers. Twenty-five topics for speaking were set internally. This situation was summarized by West and Crighton (1999) in their report on examination reform in Central and Eastern Europe, "The previous system permitted teachers to set (mostly oral) questions reflecting their own teaching rather than the full curriculum, which meant that entire content areas could be omitted without any external check" (West & Crighton, 1999:271).

## 2. Changes in ELT in Central and Eastern Europe

The active approach of the lecturers from the British Council and the American Embassy had a big impact on the teachers of English. On their own initiative, the teachers attended in-service teacher training sessions in order to improve their own language competence and to become aware of new ways to teach EFL. Communicative language teaching rapidly became very popular and teachers replaced their traditionally written textbooks by domestic authors for textbooks written by native speakers. The goal of foreign language teaching became communicative ability.

In the early stages of the intensive changes in ELT Central- and Eastern-European teachers started to see a considerable gap between teaching and assessment. Assessment did not reflect what many teachers taught in their classes putting the emphasis on immediate language practical application outside the classroom. The students' results were not comparable due to the way of storing data, and could not provide reliable data. The subjective system of assessment discouraged teachers who used modern methods in ELT. Increasing dissatisfaction led to national and later international discussions mostly in Slovenia, Slovakia, Poland and in the Baltic States. The reasons were the following:

- The examinations did not meet basic quality criteria such as reliability, validity and objectivity.
- No standardised marking schemes were provided and used.
- No external verification of standards was set.
- Basic procedures necessary for test design were not met.

The first changes occurred in Slovenia. In 1991 an attempt was made to impose an internal final examination in four subjects, which provoked student strikes and protests. In March 1992 the national *Matura* commission was established and in June 1993 the National Examinations Centre became an independent body by law. *Matura* reform (since January 1994) resulted in the first external *Matura* examination, held in June 1995. The positive effect influenced other countries like Slovakia, Hungary, Poland and Latvia.

In 1996, the second draft of the Common European Framework of Reference became a challenge for language teachers due to the original idea of the project - to facilitate mutual recognition of qualifications, and to confer on objectives and achievement standards. This motivated Slovak teachers of English and the lecturers from the British Council in Bratislava to implement a project to introduce the ideas from the CEFR into the English syllabus, and consequently into the school-leaving examination. This effort resulted in the first national piloting of the English test in March 1997.

The Lisbon Recognition Convention, signed in Lisbon on 11 April 1997, promoted the trends that supported a convergence in European education requirements. The Convention was concerned with the recognition of documented competence, knowledge and skills without recourse to repetition of assessment, examination and testing of such competence, knowledge and skills. Central and Eastern European countries started a very intensive preparation of examination reform.

The official version of the Common European Framework of Reference: learning, teaching, assessment (henceforth CEFR) was published by the Language Policy Division of the Council of Europe in 2001 after a period of research and piloting that lasted from 1993 to 1999. During a period of piloting the CEFR, modern language teaching methods changed the prevalent attitude on language learning knowledge which was based just on grammar and vocabulary. The view started to include knowing what language is appropriate for use in a given situation and how suitability differs from one culture to another. These new approaches were definitely promoted in the CEFR, which reminded language teachers to extend the narrow linguistic range of teaching by including other competencies such as pragmatic competence, sociolinguistic competence, intercultural competence, strategic competence and existential competence. In Slovakia, in a two-year period (1997-1999) new documents like the syllabus for level B1 and the syllabus for level B2 were presented at national conferences. The procedure and the new ideas based on referential levels, scales and descriptors were adopted in the Catalogues of Requirements (guidelines for students who decided to take the school-leaving exam in English).

The results achieved in the tests piloted in 1997-1999 were presented at various conferences in Europe and compared with similar approaches developed in Hungary, the Czech Republic, Slovenia, Latvia and Poland. The ideas about learning, teaching and assessment borrowed from the CEFR were influenced by various national contexts. While in Hungary a lot of effort went into including mediation into the school-leaving examination, other countries insisted on testing grammar and vocabulary in the Language in Use part.

Central- and Eastern-European language teachers quickly recognised the changes promoted by the Language Division of the Council of Europe and actively became involved in the activities by Johanna Panthier and, later, by Waldemar Martyniuk. A good example of an active approach can be seen in the papers written by Hanna Komorowska from Poland, describing the process of introducing the CEFR in pre- and in-service teacher education. Her experience resulted in recognising “the value of the CEFR” principally in positively stated levels and descriptors (can-do statements), which has become the best known part of the document (Komorowska 2004:63).

It is necessary to emphasize that many training sessions were organised all over Central and Eastern Europe aimed at familiarisation with the scales, levels, and descriptors. In Slovakia, the immediate impact of the CEFR was reflected in the new English syllabus design. The CEFR describes language use and learning as competence-based, giving an insight into the range of different competencies involved in learning. The new Slovak communicative syllabus included, communicative competencies (linguistic, sociolinguistic and pragmatic), in addition to general competencies. “Can do” statements were used in the national context and two levels (B1 and B2) were used as the reference for setting objectives. Later, other language syllabus designs appeared (for teaching German and French), and the scale was used to compare levels both within and across languages. Slovak language teachers attending international conferences could recognize the similar processes in neighbouring countries, such as Slovenia, Hungary, Poland, Latvia etc. The enthusiasm for the document was felt not only across Europe and the sheer number of translations is evidence of its power.

The CEFR set out the range of options for learning objectives, about syllabus design, classroom methodology and testing. Due to the fact that the former socialist countries were challenged to change their traditional systems of teaching and testing, many conferences were held, for instance The Central and Eastern Europe Symposium on Testing and Curriculum Design for the Final School Years in Oxford in 2001, English Language Assessment in the Central European Context in Slovakia in 2002, the IATEFL Hungary Conference in Budapest in 2003 with many presentations related to the use and value of the CEFR. The Hungarian perspective was very challenging. They were the first ones from that European region to introduce mediation, a term borrowed from the CEFR global scales.

Hungarian, Polish, Slovak and Slovenian experts of the CEFR, the most influential document of the decade, provided expertise in a workshop focused on benchmarking the ETS TOEFL scores with the Common European Framework in 2004 in Utrecht (NL). The second expertise meeting was organised in 2006 in Munich (D).

All documents associated with the CEFR influenced target language learning and teaching in the national contexts. The CEFR has been translated into 37 languages (15 of which are languages of former socialist countries or parts of these countries) and is being translated into two more (both languages were part of the post-communist area). The CEFR is associated with the European Language Portfolio, 118 (29 from the above-mentioned countries) official versions of which use checklists of CEFR-based descriptors for different levels to set individual priorities and track progress. Associated documents such as the Dutch CEFR Grid, illustrative samples of speaking performances, reading tasks, etc. have influenced the view on language teaching and testing and are reflected in many curriculum designs and language examinations.

Despite the fact that textbooks claiming to be based on levels are being published, and a wide range of internationally recognised tests labelled by CEFR-based levels are offered, both theoretical and empirical evidence of the existence of such a relation needs to be provided. The first draft of the Manual (Relating Language Examinations to the Common European Framework of Reference for Languages: Learning, Teaching, Assessment (2003) had an enormous impact on the individual claims and, before long, the first projects started to be implemented in linking language examinations to the CEFR (e.g. *Towards relating national examinations to the Common European Framework of Reference - A Hungarian example*; Benke, 2005).

### **3. The CEFR and Its Role in Language Teaching and Assessment in Central and Eastern Europe**

The CEFR has played an important role in changes that have occurred since its introduction in Central and Eastern European countries. There are many reasons why educational institutions in these countries have decided to adopt the CEFR approach to language education. As most of these countries expressed their wish to join the European Union, local governments followed a political agenda for language teaching as an instrument in the development of 'Democratic Citizenship in Europe' (the title of the Council of Europe's modern languages project from 1997 to 2001) and started to support learning languages due to the conviction that knowing different languages is a powerful factor in intellectual development. The CEFR attracted not only the language teachers but also the institutions due to the fact that it is based on communicative approaches to teaching, which include co-operation in learning, and the development of reflexive and autonomous learning habits. It also promoted methods of modern language teaching concentrating on the language learner and user both as an individual and as a social agent. This in turn promoted the idea of greater mobility, more intensive personal interaction and a deeper mutual understanding.

Competence-based teaching, learning and assessment, as well as the same descriptor scales for these three areas became popular with the language teachers, and later were officially recognised by the ministries of education in Central and Eastern European countries. In addition to teaching, assessing language competence can be considered the fundamental change. Language teachers from Central and Eastern Europe used to test only grammar and vocabulary, using translation of mother tongue sentences with particular grammatical structures into a target language. Currently competence-based teaching has resulted in the national tests that are comprised mostly of five parts: reading comprehension, listening comprehension, grammar in use, writing and speaking. The parts of the national tests are dependent on the purpose and each country's requirements. Moreover, they must meet the standards of the international language tests. The changes in primary and secondary education has become a driver for universities which have had to change their major language curricula to prepare language teachers to teach languages how they are used and to accept action-based approach, not knowledge-based.

### **4. Background**

The CEFR is a descriptive framework, which sets out a range of options, enabling specific courses or national and international examinations to be described in terms that will identify similarities to, and differences from, other courses and examinations (Morrow,

2004). This became an enormous challenge in school-leaving examination reform, therefore the new syllabi and requirements were based both on good traditions in the Slovak educational system and new trends in ELT.

At the heart of the CEFR are the Common Reference Levels (the global scale) with a broad description of what a user of a language can do ranging from basic A1- A2 through independent B1 - B2 to proficient C1 - C2. They function as reference points both for descriptions of levels, and for definitions of objectives. After a two-year study, the Ministry of Education of the Slovak Republic recognised a two-level system (B1 and B2) proper for secondary-school leavers.

The Slovak school-leaving examination reform unofficially started on 12 March 1997 when the first test in English was piloted by 1,000 students who volunteered in 40 secondary grammar schools across the country. The pilot test was written by a team of methodologists, teachers of English and native speakers - British Council teacher trainers. The new way of assessment was presented in a number of conferences throughout the country with positive effects. From 2000-2003, the pilot tests were prepared by the National Institute for Pedagogy and thousands of volunteering secondary schools were involved in it. The first pilot test in English officially recognized by the Ministry of Education was given in 2004. Since 2005 English has been tested externally with an increasing number of applicants each year.

## **5. Familiarisation**

The first stages of the linking process started in 2006 when the teachers of English were trained to become familiar with the CEFR. Later, these 32-hour courses became part of a modified project, developed by the National Institute for Education and being recognized by the Ministry of Education, the training process was offered to teachers of all foreign languages (a 36-hour course). The goal of these courses was a common understanding of the CEFR levels and the ability to relate test tasks from national tests to the levels. These courses were supplemented by various activities, for instance how to assess productive skills - writing and speaking, and using Council Europe illustration samples.

Later, the national video samples of students' oral performance at B1 and B2 were produced and attached to the official assessment as material for benchmarking. "If the result of the benchmarking process is that performance samples from the test are successfully benchmarked to the levels that were intended in designing the test, this corroborates the claim based on Specification" (Manual. 2009:20).

In 2008, the first training sessions were aimed at exercises in sorting CEFR descriptors in order to become familiar with the style of description, with the categories and above all with the levels. The second training sessions focused on calibrated samples, aiming at relating the performances to the criteria given. Participants worked in pairs, discussing their conclusions in a group and then they assessed individually, discussing their assessment. After using the calibrated performances, participants were given local samples to benchmark local performance samples to CEFR levels. It is important to use the common interpretation of the CEFR levels with the calibrated tasks and items to estimate the CEFR level of local items. It helps to link scores in the test in question to CEFR levels.

## 6. Specification

In this stage of the linking process all the forms recommended by the Manual (2009) were filled in properly. The A1 form was used as a general description of the examination and then forms A2-A8 to focus on content analysis and the relationship of content to the CEFR. A team of professionals involved in test design and test specification used A8-A10 and A19. These forms are very important for raising awareness, and it resulted in improving the quality of the following test. The items of the test were verified whether they matched the requirements based on CEFR descriptors. It was easier to verify the linking of Listening and Reading due to the precisely written descriptors of the CEFR; verifying Language in Use items was more difficult due to the lack of *English Profile*, which is scheduled to be completed in 2014.

As North (2004) states the CEFR descriptors offer a practical, accessible tool that can be used to examine content (specification), in conjunction with the CEFR Reference video and calibrated reading tasks and items to train teachers, assessors, and item writers (standardisation) and to provide criteria for ratings by trained teachers and assessors that can act as an external criterion in order to validate empirically the link to CEFR levels (external validation).

Specification forms comprised all the information about the tests. The B2 test in English consisted of 53 closed items and 27 open-ended items. The B1 test comprised 43 closed items and 17 open-ended items. The intention of the Ministry of Education is to promote language learning so that all students will take the B2-level test in the future. Three-part tests included:

- Listening Comprehension - both levels: 20 items (30 minutes)
- Language in Use - B2-level: 40 items (45 minutes), B1-level: 20 items (25 minutes)
- Reading Comprehension - both levels: 20 items (45 minutes).
- Candidates were not allowed to use any dictionaries.

The **Listening** part included all pre-recorded gaps and pauses on the cassette/CD to allow for the reading of the questions and answering of the questions, as well as pre-recorded repetitions of the source material where these were provided. The recordings (interviews, radio programme announcements and commentaries) included both monologues and dialogues. Genuine authentic material was used and every effort was made to re-capture the authenticity of the original while recording in the studio. The differentiation between the levels (B1 and B2) depended on the text and items such as the complexity and density of the language, the use of accents, the nature of the task, the speaker's voice, style, etc. The operations tested were locating and understanding specific information in a recording, understanding the overall message, extracting the relevant points to summarise the whole recording.

First task	Second task	Third task
Multiple choice	True/false/not stated	Matching
01-07	08-13	14-20

**Language in Use** consisted of three separate tasks for B2 and two tasks for B1, which had candidates working with authentic material, mostly newspaper/magazine articles.



The items were created to assess the candidates in the following areas: usage of grammatical structures, range of vocabulary, practical usage of English tenses and verb forms and word formation. The differentiation between the levels (B2 and B1) depended on the text and the items, which were chosen according to the following requirements: complexity and accuracy of language, appropriate word choice.

	Multiple choice	Word formation	Gap-filling
B2-level	21-40	41-50	51-60
B1-level	21-30	31-40	-

**Reading Comprehension** (20 items) consisted of three separate tasks in which items were based on the purposeful reading of authentic material, mostly letters and newspaper/magazine articles. The tasks involved processing operations such as locating and understanding specific information in a text, understanding the overall message of a text, deciding whether the information given is relevant to the demands for completing the task.

	Matching	T/F/evidence	Gap-filling
B2-level	61-67	68-73	74-80
B1-level	41-47	48-53	54-60

Item-writers based their expectations of what can be done in terms of the complexity of the information to be extracted, the range of points to be handled from a text and flexibility in matching reading style to task.

## 7. Standardisation of Judgements

A team of language professionals (15), trained by Dianne Wall, in a 36-hour training session of familiarization, specification and standard setting activities focused on an adequate understanding of the CEFR levels, targeted on scales, descriptors and illustrative samples related to B1 and B2 levels started by lead-in and illustration. The most important words for the B1 descriptors recognized by them were: common every day or job related topics, clear standard speech, straightforward factual texts, relevant information in everyday material, reasonable accuracy in familiar contexts, a sufficient vocabulary on topics of everyday life and good control of elementary vocabulary. As far as B2 is concerned, the participants of the training session recognised these descriptors: linguistically complex speech on both familiar and unfamiliar topics, standard dialect speech, broad active reading vocabulary, detailed information in long complex texts, a high degree of grammatical control, a good range of vocabulary. Following the first draft of the Manual (2003) was very good guidance for training sessions because the activities are clearly described and supported by clear evidence. Individual ratings were compared with pair rating, and the most important part was the whole group discussion. It revealed some issues which might have been understood differently, influenced by the national context.

The training process resulted in collecting information and giving feedback. The teachers' judgements (in black) seem to be more consistent while the statistics line reflects the real achievements of the students in listening comprehension (Figure 1). Each year the students complain about the task and item difficulty in the listening section. The statistics show that the techniques used in listening comprehension section are deployed. Students were able to solve the problems they were exposed to and achieved scores are higher than expected.



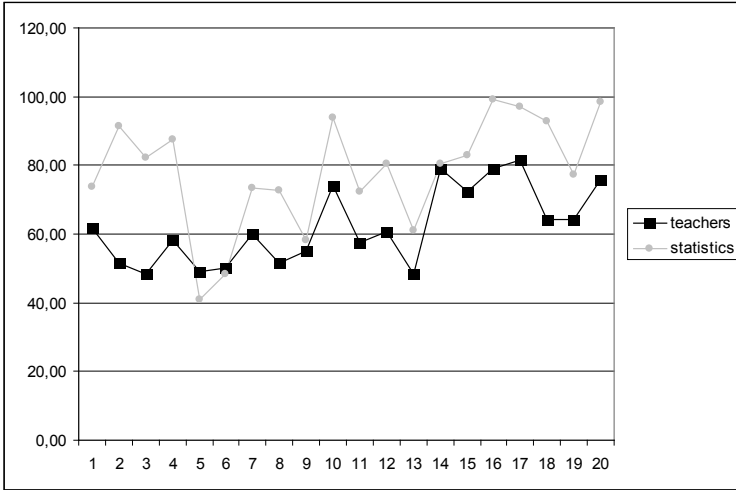


Figure 1 Listening Comprehension: Teachers' judgements and statistics

The differences between the teachers' judgements and statistics in a multiple choice task (01-20) are shown in Figure 2. The statistics proved that distractors were highly effective in differentiating good students from weak students. Word formation tasks comprise items 21-30. Gap-filling tasks are included in items 31-40. Due to teaching being still based a lot on grammar, the teachers' expectations were higher than the actual students' scores. The tests were designed to measure the level of grammar and vocabulary acquisition. Many professionals who use the CEFR and highly appreciate that it encourages its users to look at the skills as being integrated rather than isolated are a bit disappointed by the lack of grammar patterns and vocabulary lists for each level. Some countries presented profiles (German Language Profile) in which such structures are proposed.

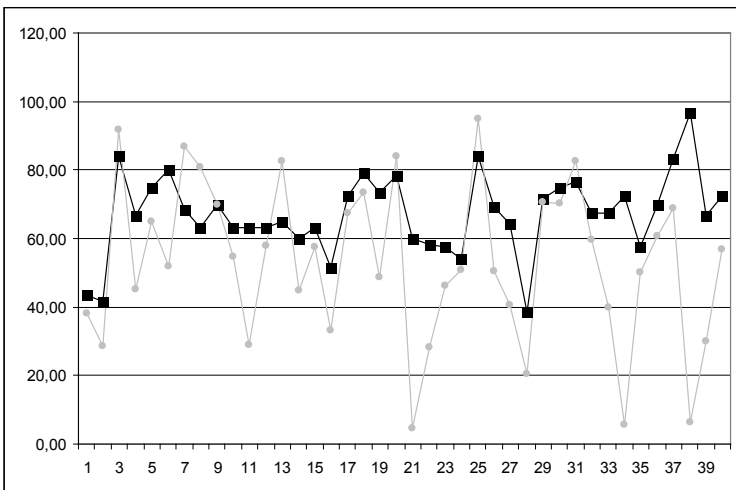


Figure 2 Language in Use: Teachers' judgements and statistics

The teachers' judgements of reading skills (Figure 3) are more consistent than actual performances. Due to their experience that many students tend to read word by word, to lose the meaning of sentences as soon as they are decoded, not to skip words and to lose a lot of time hesitating, when they encounter new words, the teachers expected poorer students' performances.

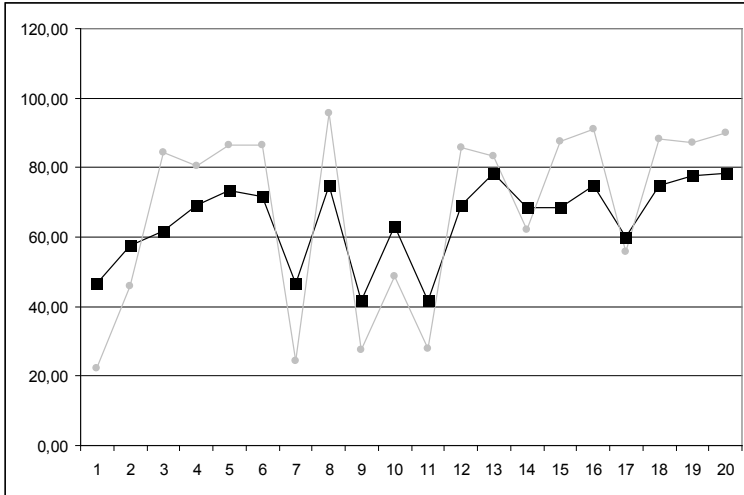


Figure 3 Reading Comprehension: Teachers' judgements (in black) and statistics (in grey)

To sum up, comparing the teachers' judgements and students' scores it is possible to state that feedback from teaching in class influences teachers' judgements. According to the results of the questionnaires, the teachers still spend a lot of time practicing grammar and vocabulary, and therefore they expected the same score in receptive skills as in Language in Use. Teachers' judgements for Listening Comprehension was 66.6 and the official statistics was 78.4, Language in Use 66.78 in teachers' judgements and 54.6 in statistics, and Reading Comprehension was 64.75 in teachers' judgements and statistically 78.7. Listening (66.6) and reading (64.75) judged by teachers are quite close. The similar closeness is seen between 78.4 in listening and 78.7 in reading provided by the statistics. The teachers' judgements for language in use were closer to receptive skills judgements than the official statistical data.

The histogram (Figure 4) shows that the distribution of scores can be considered as normal (bell-shaped). Some basic statistics for the B2 test are as follow: the number of students tested – 6.753, the mean – 70.5%, the median – 71.7%, the mode – 71.7%, the standard deviation – 12.3. Cronbach's alpha for B2 was .88 and for B1, .91. The data and the histogram suggest that the test was not very difficult for the tested population. The test differentiated the successful students from less successful and weak students.

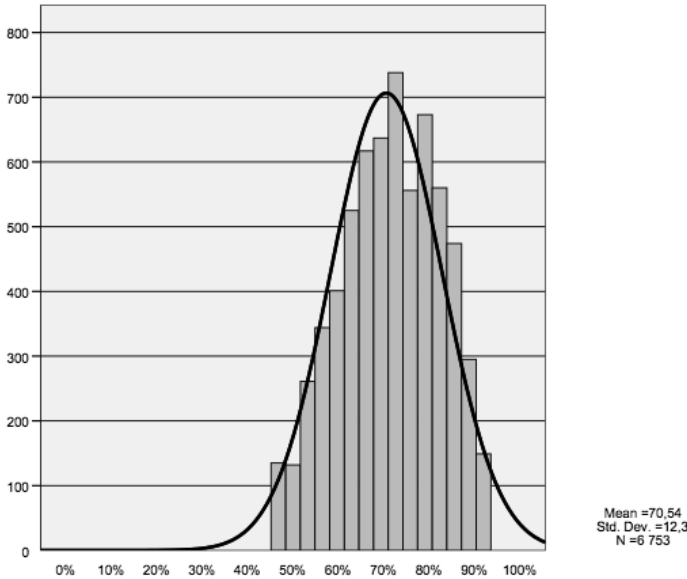


Figure 4 Distribution of students' B2 test results

The coefficient of correlation ( $r$ ) is a measure of the relationship between two measures. A correlation coefficient value of exactly +1 indicates a perfect positive fit. A correlation greater than 0.8 is generally described as strong, whereas a correlation less than 0.5 is generally described as weak. The correlation between the students' Language in Use scores and their total scores was .86. A correlation coefficient value of exactly +1 indicates a perfect positive fit. Table 1 illustrates the correlation between the students' Language in Use scores and their total scores.

**Correlations**

		Red. Score - Grammar	Total Score
Red. Score - Grammar	Pearson Correlation	1	,862**
	Sig. (2-tailed)		,000
	N	6753	6753
Total Score	Pearson Correlation	,862**	1
	Sig. (2-tailed)	,000	
	N	6753	6753

\*\* . Correlation is significant at the 0.01 level (2-tailed).

Table 1 Correlation between raw scores from the B2 Language in Use part and the B2 test

The co-efficient of correlation between the teachers' judgements and official statistics in the B1 test was 0.71, higher than in the B2 test ( $r = 0.55$ ). Likewise, for B1, as the values for Language in Use increase, values for total scores also increase ( $r = 0.742$ ). Positive values indicate a relationship between the given two variables (the students'

performance in Language in Use and their total scores) such as illustrated in Figure 5 (as values for Language in Use increase, values for total score also increase).

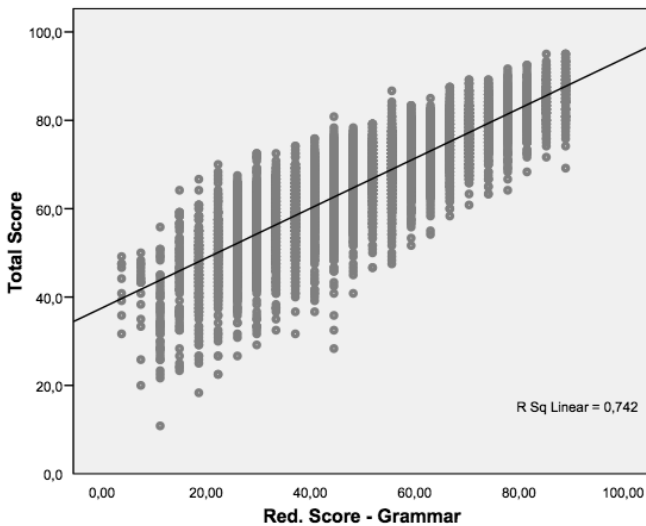


Figure 5 Correlation between standard scores from B1 Language in Use part and the B1 test

To conclude, the ability of students to use language in use in both B1 and B2 tests was significantly related to their total scores in both tests. On the one hand, the teachers would judge their students poorer than their actual performances in reading and listening; on the other hand, they overrated them in language in use. Interpreting the score in a given test as an indicator of an individual's ability, that score must be both reliable and valid, as shown here. In order to link the examination to the CEFR, it is recommended to develop, apply and report transparent, practical procedures in a cumulative process of continuing improvement. The tendency to relate these tests to the CEFR was strong as many necessary procedures were followed such as describing the examination coverage, analysing the quality of the examination and providing supporting evidence that reports the procedures followed.

## 8. Linking the School-leaving Test in English to the CEFR

The linking process uncovered a number of issues associated with the CEFR. Teachers, teacher trainers, test designers and item writers need regular training sessions which will allow them to discover more and more useful material covered in the document. Language professionals should be trained how to use this document with the fruitful support of supplementary materials made public on the website of the Council of Europe. The data presented in the statistics were reliable and valid. All examination providers have some form of standardisation training with performance samples for productive skills, used to train examiners in reasonably consistent application of the assessment criteria, or standard-setting procedures for receptive skills, used to ensure a reasonably consistent level of difficulty in the tasks and items set.

Research followed classical procedures of quality. All the stages had a positive impact on the project participants who started to view previously assumed issues from different viewpoints. The process was notoriously difficult. Initial estimates of the level of difficulty of a task bore a limited relationship to the actual difficulty. But in any national level assessment, empirical validation is a requirement. For internal validation, techniques and procedures of Classical Test Theory (CTT) were used. It is also necessary to include Item Response Theory (IRT) which focuses on the concept or construct to be measured. IRT methods require the data model be checked and its validity demonstrated. External validation concerned the verification of the standard-setting undertaken in relation to CEFR levels. Research was based on correlation between the scores on the examination and teachers' judgements. This is a traditional way of demonstrating the construct validity. The number of procedures involved in the validation of a claim that the school-leaving test in English is linked to the CEFR is required. The report on empirical validation will have to provide sufficient information and references to enable a critical analysis of the procedures and methodologies used.

## 9. Conclusion

The CEFR is becoming more and more popular in the countries of Central and Eastern Europe. The document is still topic number one in conferences and most projects related to teaching languages focus on associating the institutional documents with the CEFR. The project implemented by Masaryk University in Brno, the Czech Republic, is a good example. The idea of relating the official assessment tool at universities - the European Credit Transfer System grading scale into the three-level descriptors (B1, B2 and C1) based assessment grid is still being piloted.

So far the CEFR has influenced teaching and testing in most Central and Eastern European countries to a great degree and many national pedagogical documents are claimed to be associated with it. It has influenced material writing, the syllabus or curricula design as well as test development. As one of the most important educational documents of the Council of Europe, the CEFR is mentioned in pre-service or in-service teacher training at universities in Central and Eastern Europe.

Its importance is evident in a number of projects implemented within the European Social Fund. In Slovakia, the National Institute for Education offered elementary school teachers an eight-semester course based on the national curriculum to acquire the target languages at Level B2. The system is based on regularly testing their competences at each level. The levels are being used as the reference for setting objectives.

As far as the countries of Central and Eastern Europe are concerned, the traditional system of teaching in some parts, a lack of qualified language teachers and dissatisfaction with foreign language competence in the countries make the governments think about ways to improve acquisition of the target languages. The countries of Central and Eastern Europe are developing their language tests with the goal to demonstrate the level of language competence of their primary and secondary school leavers and university graduates. Several countries started the process of relating their final examinations to the CEFR. The procedure is not very easy and needs a lot of effort and a serious approach to validate the claim. Many tests are designed pretending to be based on the levels without any evidence.

The *Common European Framework of Reference: learning, teaching, assessment* is beneficial for the quality of teaching/learning target languages as well as testing language competence. Linking the institutional documents and national examinations to the CEFR is expected to initiate strategic actions within three major areas in education: standards, evaluation and professional development.

## References

- Bachman, L. F. 1995. *Fundamental Considerations in Language Testing*. Oxford : Oxford University Press.
- Bachman, L. F., Palmer, A. S. 1996. *Language Testing in Practice*. Oxford : Oxford University Press.
- Benke, E. 2005. *Towards Relating National Examinations to the Common European Framework of Reference (A Hungarian example)*. [www.nyelvvizsgak.hu](http://www.nyelvvizsgak.hu) [Retrieved on March 29, 2011].
- Bérešová, J. 2004. *Komunikatívne testovanie cudzích jazykov v súlade s požiadavkami Rady Európy*. Trnava : Trnavská univerzita.
- Council of Europe. 2001. *Common European Framework of Reference: Learning, Teaching, Assessment*. Oxford: Oxford University Press.
- Komorowska, H. 2004. The CEF in pre- and in-service teacher education. In: Morrow, K. *Insights from the Common European Framework*. Oxford: Oxford University Press, pp. 55-64.
- Manual. Relating Language Examinations to the Common European Framework of Reference for Languages: Learning, Teaching, Assessment*, 2009. Strasbourg: Council of Europe Publishing.
- Morrow, K. 2004. *Insights from the Common European Framework*. Oxford : Oxford University Press.
- North, B. 2004. Relating assessments, examinations, and courses to the CEF. In: Morrow, K. *Insights from the Common European Framework*. Oxford: Oxford University Press, pp. 77-90.
- North B. 2005. *The CEFR Levels and Descriptor Scales*. Paper presented at the 2<sup>nd</sup> ALTE International Conference, Berlin 19-21 May 2005.
- West, J., Crighton, J. 1999. Examination Reform in Central and Eastern Europe: Issues and Trends. *Assessment in Education*, Vol. 6.