The Common European Framework of Reference for Languages: Insights for language testing

Paşa Tevfik Cephe [a] *, Tuğba Elif Toprak [b]

[a] Gazi University Department of Foreign Language Education, Ankara 06500, Turkey
[b] Gazi University Department of Foreign Language Education, Ankara 06500, Turkey

Abstract

The *Common European Framework of Reference for Languages: Learning, Teaching, Assessment* (Council of Europe, 2001) which aims to bring about harmonisation and transparency within Europe and promote the idea of European citizenship is a project that provides the stakeholders with a reference document that could be utilised not only for developing language curriculum and syllabus, preparing course books but also for evaluating the learning outcomes. In terms of language testing, the major claim of the CEFR is its potential to be used as a reference point to design of new language tests and make a comparison among the existing language tests by setting standards. Council of Europe (2001) also makes clear that the CEFR could be used for the specification of the content of the test and exams, setting the criteria of assessment and describing the levels of proficiency in tests. Though the CEFR has a great potential for playing a crucial role in language testing, the issues of developing and aligning tests to the CEFR need to be considered with a critical eye. Hence, the present study is an attempt to examine the practical considerations and potential problems related to the CEFR in terms of language testing and to discuss some practical implications for language testers and language teachers in terms of test generation and alignment.

Keywords: CEFR; language testing; test development; test alignment

1. Introduction

As noted by Figueras et al. (2005), The Common European Framework of Reference for Languages: Learning, Teaching, Assessment (Council of Europe, 2001) (henceforth, the CEFR) has been one of the most influential and powerful developments in the domains of language teaching and testing in the last decade. The CEFR is a project of Council of Europe, which fosters harmonisation and transparency among cross-national institutions and promotes European citizenship (Fulcher, 2004). Examined from the dimension of foreign language testing, it is alleged that the CEFR enables language testers to generate tests with common principles that are in accordance with the values of the Council and the idea of European citizenship. According to North (2007), in that vein, the CEFR aims to create a shared meta-language that could be used to talk about aims and assessment, to stimulate

---

*Paşa Tevfik Cephe Tel.: +90-312-202-8476
E-mail address: pcephe@gazi.edu.tr

practitioners to think about their practices by considering their context, and to reach a consensus on common reference points. Council of Europe (2001) states that the main goal of the CEFR is to foster reflection, communication, discussion among the practitioners in the domains of language teaching and assessment. Besides, it was intended that the CEFR would create a basis for mutual recognition of practices in language teaching and assessment all round the Europe. This claim, regarding the use of the CEFR in language testing made by Council of Europe, is defined clearly in the following uses:

1) for the specification of the content of tests and examinations;
2) for stating the criteria to determine the attainment of a learning objective;
3) for describing the levels of proficiency in existing tests and examinations thus enabling comparisons to be made across different systems of qualifications. (Council of Europe, 2001, p. 178)

Fulcher (2004) discusses the role of the CEFR in language testing and explains the emergence of the CEFR with two broad reasons; the introduction of the European Language Portfolio which fosters goal setting and self-assessment in foreign language learning and the need to provide stakeholders with a means to compare existing tests that could enjoy recognition all round the Europe. Similarly, Bechger, Kuijper and Maris (2009) suggest that providing a descriptive system of language activities involving different levels of proficiency could be used for existing tests and examinations that are being developed. This is the point where the CEFR comes into play. The CEFR includes descriptive scales, the most general one being the global scale of common references, which labels learners as basic user (Level A), independent user (Level B) and proficient user (Level C). A further distinction is made between these levels as A1, A2, B1, B2, and C1, C2. Can-do statements, the illustrative descriptors are provided for each sub-level across different language skills namely listening, reading, spoken interaction, spoken production, and writing. Apart from behavioural aspects, the CEFR is claimed to provide stakeholders with the qualitative aspects of spoken language such as range, accuracy, fluency and interaction (Council of Europe, 2001) and to pay attention to elements such as sociolinguistic appropriateness, flexibility, turn taking, coherence and cohesion.

The CEFR has been embraced by practitioners, institutions at national and international levels so enthusiastically that curricula based on the CEFR have been developed and course books intended to help realise the goals of the CEFR have been written. Moreover, when it comes to the domain of language testing, claims about tests measuring language ability at a level intended on the CEFR have been made. Though it is stated by Council of Europe (2001) that the CEFR could be used as a reference point in both comparing existing tests and developing new tests, scholars in the domain of language testing seem to take a stance against the use of the CEFR in language testing and question prevailing practices for this aim. This paper aims to review the existing notions and discussions in language testing related to the use of the CEFR in the processes of alignment and development of language tests and proposes a model that could be used for above mentioned purposes

2. The CEFR and its use in language testing

As noted above, for the aim of using the CEFR in language testing, several steps have been taken by Council of Europe (2003; 2009) the most notable example being the manual prepared by Figueras et al. (2005). In spite of these efforts, increasing criticism has been targeted at the use of the CEFR in language testing. These criticisms come from two fronts, attacking the theoretical basis of the CEFR, specifically questioning the notion of validity and practical issues such as test content, context, rating process and so forth.

Related to the arguments about its theoretical aspects, Fulcher (2004) describes the CEFR as purely descriptive and claims that the distinction between Waystage and Threshold is not drawn by basing on

any empirical evidence, but merely on the intuition of the developers. He further attracts attention to the issue that although the CEFR includes situations, competencies, functions and abilities, it is not clear at what stage a learner would perform in above mentioned elements and what is needed to be an individual at the Threshold level. Fulcher (2004) concludes that this situation may be the result of the development process of the CEFR, as described in North (2000a) in detail.

The development of the CEFR was comprised of four phases, which were intuitive, qualitative, quantitative analyses and replication phases. Fulcher (2004) argues that it was the teachers' perception, not the learner competency that was incorporated into the scales. As North (2000, p. 573) himself puts it "what is being scaled is not necessarily learner proficiency, but teacher/raters' perception of that proficiency— their common framework". This perception, as criticised by Fulcher (2004), may be deemed as the perception of European teachers and testers and the term common refers to the agreement among them. Alderson (2007) also cautiously warns that this perception, offered by language teachers who are not trained to be testers, may not produce satisfactory results.

Another criticism directed at the CEFR is about its nature, related to the question of whether the CEFR is a framework or not. From the viewpoint of several scholars (e.g., Milanovic, 2002; Fulcher, 2004), the CEFR is regarded as operating at an abstract level like a model, rather than a true framework where content and test specifications are clearly defined. To be more specific, the CEFR is seen as a model reflecting the theories of communicative language testing as proposed by Bachman (1990). A true framework, according to Weir (2005), should help stakeholders to discover both processing and contextual elements and the relationships between them at different proficiency levels. The fact that several researchers (Huhta et al., 2002; Jones, 2002; Alderson et al., 2004; Morrow, 2004) have had difficulties in aligning tests to the CEFR, in Weir's view, can be linked to the CEFR's deficiencies at both fronts.

A different view related to the nature of the CEFR has been proposed by Alderson et al. (2004; 2009) in which the CEFR has been described as a theory of language development since it is predominantly involved in describing language use. However, Alderson et al. (2004; 2009) cautiously add that can-do statements are largely related with behaviours rather than reflecting a theory of development. Moreover, they question whether can-do statements may be converted to the items exemplifying different proficiency levels specified in the CEFR. Fulcher (2004) draws attention to the danger of the belief that the scales presented in the CEFR reflect a theory of language development, and warns that this belief may be held by the teachers. Fulcher (2004) further argues that, at the institution level, this belief even may pose greater problems since many testing agencies and test developers may claim links between the scores and the CEFR levels for the sake of getting international recognition. Apart from the issue that the CEFR cannot be regarded as a theory of development, the obscurity of constructs and their definitions may lead to a chaos in aligning studies.

Alderson (2007) commenting on the problems likely to be encountered during alignment studies states that his team had difficulties while working on DIALANG project (applying the CEFR to diagnostic testing in 14 languages) in terms of terminology and the theory of language development. Alderson (2007) seems very cautious about the methodology of the development of the CEFR and adds that although the CEFR assumes that a communicative activity requires a certain proficiency level in any language, this assumption has not been validated by empirical research. Further, he questions whether the development of the CEFR is based on second language acquisition (SLA) research. Considering the fact that majority of the research conducted within the domain of SLA focus on English, the empirical studies dealing with other languages seem to be necessary. While developing the Dutch CEFR Construct Project, Alderson et al. (2006) examined whether the CEFR was used to generate reading and listening tests and concluded that the CEFR displayed problems in terms of

clarity of definitions, overlaps and consistency. Besides, Alderson et al. (2006) argued that scales do not represent a clear theory of language development not only to diagnose, but also to test the language skills. Hence, Alderson (2007), by taking above mentioned issues into consideration, calls our attention to the fact that testers', publishers', book writers' and teachers' pure faith in the CEFR might have negative effects since these claims are not validated by empirical research. He further alleges that politicians and civil servants with no expertise in language learning and teaching also attempt to set standards, which could be harmful.

Another prominent figure in language testing, Weir (2005) also criticises the theoretical and practical foundations of the CEFR by putting forward several arguments. Though the CEFR has been described as being comprehensive, coherent or transparent for uncritical use in language testing (Council of Europe, 2001), Weir (2005) claims that the scales are based on contextual variables-performance conditions that are not complete. Weir (2005) proposes a notion of internal validity that is comprised of three units as context validity, theory based validity and scoring validity. Weir (2005) uses the term 'context validity' to refer to the social features of a task such as the setting and linguistic and social requirements. Taking different proficiency levels of the CEFR into consideration, it could be said that the participants will have to deal with various contextual conditions while carrying out a task and test developers need to pay specific attention to the constructs and contextual variables that influence test performance. These contextual variables include purpose, response format and time constraints and demands of the task.

The second criticism raised by Weir (2005) aiming at the CEFR is related with the issue labelled as 'theory based validity' that is related with the cognitive processing that examinees carry out while dealing with the tasks. However, it is alleged that the CEFR does not equip language educators with necessary views on cognitive processing at any level. Thus, it could be deduced that the scales in the CEFR do not represent an acquisitional hierarchy; an issue that has been addressed by several scholars as well (Fulcher, 2004; Alderson, 2007). The third dimension of validity in view of Weir (2005) is 'scoring validity', which is assumed to be related with the quality of performance. Weir (2005) contends that knowing how successfully an examinee should perform on a task at a specified level is necessary and this knowledge should be elaborated in terms of context based and theory based dimensions of the construct in question. Apart from scoring criteria, qualities of test raters and rating process should be taken into consideration. According to Weir (2005) scoring validity is the issue on which the CEFR has almost nothing to offer.

Apart from the theoretical considerations, several criticisms have been made regarding the practical issue in utilising the CEFR for test development and alignment issues. For instance, choices related with content choice are left to the test developers. Draft developed by Council of Europe (2003) does not elaborate on decisions about the content. Another significant problem raised in the relevant literature is about the wording of the can-do statements, which act as specifications while devising, and aligning tests (Weir, 2005; Jones, 2002; Alderson et al., 2004). Difficulties were experienced in aligning studies when the researchers recognised that there were cases in which statements were not distinguished from the levels below. This problem may have to do with the specification of context in which a task is carried out. Weir (2005) and Alderson et al. (2004) propose that the purposes for which we utilise language at different levels and context are crucial. To illustrate, it could be said that the type of reading activity that would be carried out will be based on the purpose of reading. Hence, as suggested by these scholars if we again take reading comprehension as an example, the subskills of comprehension that make up reading comprehension construct and types of reading should be taken into account while designing test task. It could be said that defining the specifications of a construct cannot be regarded as less significant than proving its statistical rigour. Alderson et al. (2004) for

instance, while working on the Dutch CEFR Construct project encountered several problems related with the test specification in terms of expressions used in can-do statements. To illustrate, they identified that eight different verbs referred to comprehension and these verbs were scan, monitor, understand, select, obtain, evaluate, locate and identify in B2 level. Alderson et al. (2004) were curious about whether this situation was related with stylistic synonyms or reflected differences in terms of cognitive processing. To overcome this problem, they decided to resort to theories of comprehension. Weir (2005) proposes that this may have to with the issue that the CEFR has a clear sociolinguistic focus but on the other hand, it attaches little attention to the psycholinguistic issues. He further advocates that to ensure a theory based validity, underlying mechanisms should be understood and should be paid specific attention.

There are several practical issues related with the CEFR mentioned in the relevant literature, mainly on test method. Weir (2005), for instance, touches upon the response format since The CEFR does not include much detailed account on test method. However, it should be noted that test format has considerable effects on the context and processing. At this point, it could be said that a reading test utilising multiple-choice questions and another reading test using open-ended comprehension questions differ a lot in their nature and also the processes they invoke. Another issue related to test method is time limit to be set. Test developers should arrange a time limit by taking the amount of time needed to carry out a specific task into account. If not, this could result in construct underrepresentation. The third issue related to test method is related with genre, discourse types and their suitability across different language levels. Weir (2005) and Huhta et al. (2002) conclude that the CEFR does not provide sufficient guidance on this point. As Alderson et al. (2004) points out, length of a test is an issue that is left vague in the CEFR as well. Topic choice also receives criticism. The CEFR neither specifies test topics nor associates topics to different proficiency levels. Since test takers general background may have an effect on their test performance, this situation should receive specific attention.

3. Aligning language tests with the CEFR

Harsch and Rupp (2011) attract our attention to the difficulty and lack of consensus on the issue of alignment and state that although the CEFR has been acting as a framework for developing language tests, in reality, it is not a manual to be used for these purposes. As a consequence, complications may arise at the point of generating tests in alignment with the CEFR. Besides, linking practices which should be taken seriously, do not entail reliability analysis, they lack theoretical background and are conducted largely on intuition. Harch and Rupp (2011) also state that language testers are not sure about how to align tests with the CEFR in terms of both practice and theory.

As a solution, North (2000b) proposes the notion of 'social moderation' in which a shared understanding of standards are determined by a group of raters through discussions and training and this process is seen as a way to link tests to the CEFR. Figueras et al. (2005) also prepared a manual for linking examinations to the CEFR in which they divided linking process into four phases as familiarization, specification, standardization, empirical validation. Familiarization refers to the activities that ensure the participants in the linking process are familiar with the CEFR. If it is detected that the participants do not have a sufficient knowledge about the CEFR, the quality of the linking process is suspected. In the second phase, specification, there is a matching procedure between categories of the CEFR and content and task types presented in the exams. If it is revealed that examinations cannot be described in terms of the CEFR categories then the alignment process becomes susceptible. The third phase is standardization in which benchmarks for a test are determined by a group of experts in accordance with the constructs described in the CEFR. The last phase is the

phase of empirical validation to ensure that both the exam and the alignment to the CEFR is well-grounded. The evidence is gathered through analysis of test data and ratings from the assessment.

4. An informed approach for language testing practices

It could be stated that the CEFR affects ongoing practices both in language teaching and language testing and it is highly likely that it would continue doing so in the future. Though it is apparent that the CEFR has several shortcomings when examined from the angle of language testing, it would be unfair to say that it has no benefits to offer. Indeed, if planned carefully from the very start by taking above mentioned issues into consideration and enriching the test design and construction process through feedback from the stakeholders, our testing practices may prove useful. The CEFR, at this point, can be a good starting point. Moreover, taking the fact that curriculum and course books are designed on the basis of the CEFR's tenets, it would not be surprising that testing would be aligned with the CEFR somehow. Below (see Figure 1.), a model incorporating the phases of testing practices and reflecting the relationships between these phases has been presented.
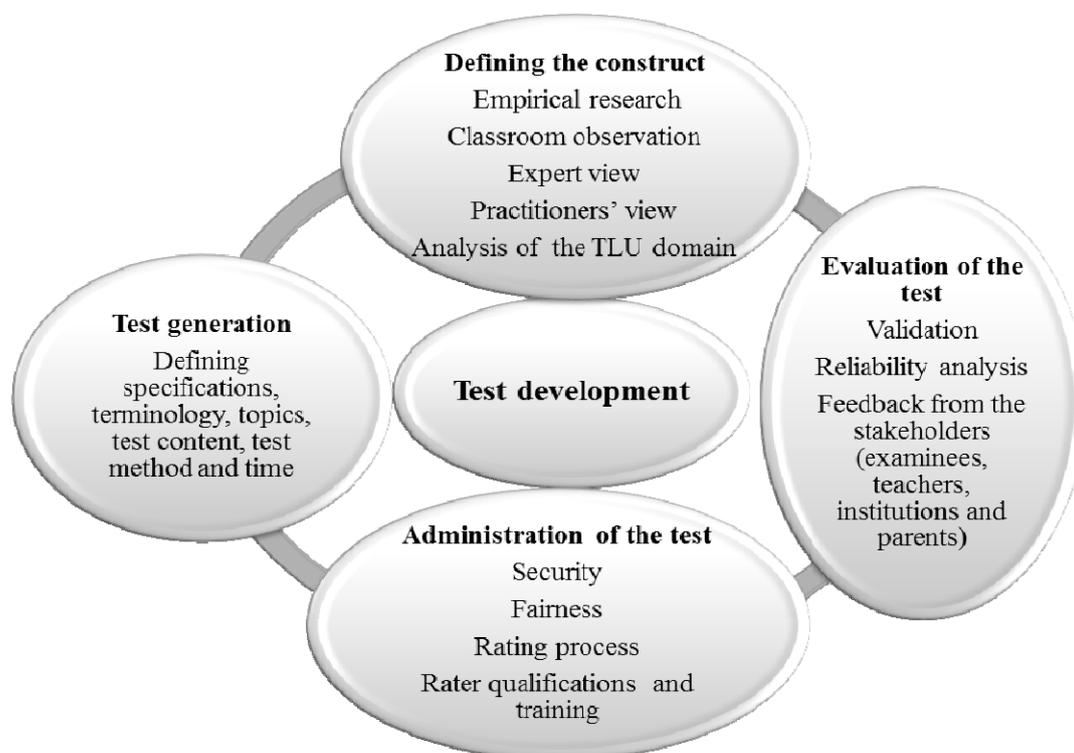
Figure 1. A model reflecting the phases of testing practices and the relationships among these phases

As can be seen in Figure 1, test development is comprised of four steps reflecting a cyclical nature. Test development is seen as a continuous process in which a step feeds the following step. In the first step, which is 'defining the construct' test developers define the construct that they intend to measure in detail and examine its properties. It could be stated that, this step is the most crucial step since test development cannot be built on ill-defined constructs. The knowledge necessary to define the construct comes from several sources such as the relevant literature, the views of experts, field observations, the views of teachers and the analysis of target language domain. Target language use

domain may be deemed of being utmost importance since each context has its unique conditions, necessities and requirements. For instance, the kind of reading activities carried out in a classroom where English is taught for academic purposes and the students are trained to be professionals would be quite different from the classroom in which English is taught for general purposes. For no doubt, this situation would affect our choices in testing practices. We could expect that addressing above mentioned issues would help us deal with the shortcomings of the CEFR in terms of the theory of language development and 'theory based validity'.

After the construct has been defined and elaborated, the test generation phase, the phase in which the test specifications constituting the basis of the tests are formed. This is the phase where test developers engage in test construction work and design the architecture of the test. During test generation, issues apart from the specifications, such as the use of terminology, topic selection, test content, task type and duration are also dealt with. Since each of these factors may shape the performance of the test takers, careful analysis, detailed planning and informed approach are needed before we make our choices. Given that the CEFR does not say much about the test method and for no doubt an elaboration on this issue is needed, precise decisions we make regarding the test construction and application would help us overcome the shortcomings of the CEFR in terms of 'context and scoring validity'.

In the third phase, the test is administrated by taking several points into consideration; and these topics would be secure and fair application of the test, rating process and the qualifications of the raters assessing the performance of the examinees. Since test scores are used to make decisions about the educational and professional lives of the test takers, and therefore are highly likely to affect their lives to a great extent, the fair and secure administration of the tests should be of utmost importance so that undesirable outcomes would not occur. Assuring that the test is applied fairly and securely would not be enough on its own since rating process and qualifications of the people carrying out rating task are also vital. It must be ensured that the raters possess an informed approach and be equipped with the necessary knowledge about both the domain and the skills that they are assessing, and also the tenets of evaluation practices. At this point, it would be very essential that the raters should be trained beforehand to cope with these issues.

The last phase, the evaluation phase is the step in which the interpretations arising from test scores are validated, reliability analyses are conducted and feedback from the stakeholders are received. Evaluation phase is of utmost importance since a comprehensive evaluation both the test and its consequences are examined. Insights gained from this step are crucial since they may be used to better our understanding of the constructs, test design and application. At this point, apart from proving the statistical rigour of the test, the evidence that would shed light on and strengthen the interpretations of the scores obtained in the test is needed to ensure the notion of validity. Moreover, since test scores affect the lives and decisions of many parties such as examinees, teachers, school boards, institutions and parents their feedback should also obtained to be used both to evaluate and design our testing.

With clearly defined specifications, an informed approach about the domain we are testing, views obtained from all the stakeholders, carefully selected test content and test method and qualified people involved in the language testing process the use of the CEFR as a reference point would be likely to yield favourable results. Especially considering the fact that language curricula, course books and instructional activities are designed in accordance with the CEFR, it could be stated that making the most of the CEFR by paying attention to its several drawbacks in terms of theory based, context based and scoring based validity and finding solutions for these drawbacks would be, in a way, building a bridge between our teaching and testing practices.

5. Conclusions

Weir (2005) contends that since the CEFR was not designed specifically for language testing, it takes laborious research, reflective test development practices to make use of it. He acknowledges that the earlier work (Threshold Level, Waystage and Vantage studies) and North's (2002) endeavours to calibre functions on a common scale the emphasis on functional competence are the strengths of the CEFR. The CEFR can be used to determine the objectives for teaching and assessment but inadequacies also prevail. Weir (2005) sees the CEFR as heuristic rather than prescriptive in nature and by taking the deficiencies in terms of validity into account, he proposes that making comparisons based only on the scales may be somehow misleading. He further alleges that the CEFR, at present, does not help us develop comparable tests let alone helping us to decide if these tests are comparable. Fulcher (2004) also argues that the CEFR may be of use in language testing as a user-oriented scale which serves as understandable, practical reporting instrument for stakeholders. Though problems have been encountered while implementing the CEFR to language testing, using the CEFR scales for reporting what a learner can do with a score in specific domain may be useful. To fight these shortcomings, a test development model which would help and guide us in defining the constructs, generating the test, administrating and evaluating it by taking our context into consideration would have clear and practical implications.

References

Alderson, J. C. (2002). *Common European Framework of Reference for Languages: Learning, teaching, assessment*. Strasbourg: Council of Europe.

Alderson, J. C., Figueras, N., Kuijper, H., Nold, G., Takala, S., & Tardieu, C. (2004). *The development of specifications for item development and classification within the Common European Framework of Reference for Languages: Learning, teaching, assessment. Reading and listening.* Final report of the Dutch CEF construct project. Unpublished document.

Alderson, J. C., Figueras, N., Kuijper, H., Nold, G., Takala, S., & Tardieu, C. (2006). Analysing tests of reading and listening in relation to the Common European Framework of Reference: The experience of the Dutch CEFR construct project. *Language Assessment Quarterly, 3*(1), 3-30.

Alderson, C. J. (2007). The CEFR and the need for more research. *The Modern Language Journal, 91*(4), 659-663.

Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.

Bechger,T. , M., Kuijper, H., & Maris, G. (2009). Standard setting in relation to the common European framework of reference for languages: The case of the state examination of Dutch as a second language. *Language Assessment Quarterly, 6*(2), 126-150.

Council of Europe. (2001). *Common European Framework of Reference for language learning and teaching*. Cambridge, UK: Cambridge University Press.

Council of Europe. (2003). *Relating Language Examinations to the Common European Framework of Reference for Languages: Learning, teaching, assessment*. Strasbourg: Council of Europe.

Council of Europe (2009). *Manual for relating language examinations to the Common European Framework of Reference for Languages (CEFR).* Strasbourg, France: Council of Europe. Retrieved from http://www.coe.int/t/dg4/linguistic/Manuel1_EN.asp

Figueras, N., North, B., Takala, S., Verhelst, N., & Van Avermaet, P. (2005). Relating examinations to the Common European Framework: A manual. *Language Testing, 22*, 262-279.

Fulcher, G. (2004). Deluded by artifices? The Common European Framework and harmonization. *Language Assessment Quarterly, 1*(4), 253-26.

Harsch, C. & Rupp, A. A. (2011): Designing and scaling level- specific writing tasks in alignment with the CEFR: A test-centered approach. *Language Assessment Quarterly, 8*(1), 1-33.

Huhta, A., Luoma, S., Oscarson, M., Sajavaara, K., Takala, S. & Teasdale, A. (2002). DIALANG: A diagnostic language assessment system for learners. In J. C. Alderson (Ed.), *Common European framework of reference for languages: Learning, teaching, assessment. Case studies* (pp. 130-145). Strasbourg: Council of Europe.

Jones, N. (2002). Relating the ALTE framework to the Common European Framework of Reference. In J. C. Alderson (Ed.), *Common European Framework of Reference for Languages: Learning, teaching, assessment. Case studies* (pp. 167-183). Strasbourg: Council of Europe.

Milanovic, M. (2002). *Common European Framework of Reference for Languages: Learning, teaching, assessment: Language examining and test development.* Strasbourg, France: Council of Europe.

Morrow, K. (Ed.). (2004). *Insights from the Common European Framework.* Oxford: Oxford University Press.

North, B. (2000a). *The development of a common framework scale of language proficiency*. New York: Peter Lang.

North, B. (2000b). Linking language assessments: An example in a low stakes context. *System 28,* 555-577.

North, B. (2002). A CEF-based self assessment tool for university entrance. In J. C. Alderson (Ed.), *Common European Framework of Reference for Languages: Learning, teaching, assessment. Case studies* (pp. 146-166). Strasbourg: Council of Europe.

North, B. (2007). The CEFR illustrative descriptor scales. *The Modern Language Journal, 91*(4), 656-659.

Weir, C. J. (2005). Limitations of the Common European Framework for developing comparable examinations and tests. *Language Testing, 22,* 281-300.

Avrupa Dilleri Öğretimi Ortak Çerçeve Programı ve yabancı dil öğretiminde ölçmeye dair görüşler

Öz

Avrupa ülkeleri arasında uyumu ve şeffaflığı sağlamayı ve Avrupa vatandaşlığı kavramını desteklemeyi amaçlayan Avrupa Dilleri Öğretimi Ortak Çerçeve Programı (Avrupa Konseyi, 2001) paydaşlara dil öğretim programı ve müfredat hazırlama, ders kitabı yazma gibi konuların yanı sıra öğrenme çıktılarını değerlendirmeye yönelik de olanaklar sağlayan bir projedir. Dilde ölçme değerlendirme açısından bakıldığında, başvuru metninin yeni sınavlar hazırlamada ve standartlar belirleyerek mevcut sınavları karşılaştırmada bir dayanak noktası olarak kullanılabileceği iddiası söz konusudur. Avrupa Konseyi (2001) de başvuru metninin sınav içeriğinin belirlemesi, değerlendirme ölçütlerinin ortaya konulması ve sınavlardaki yeterlilik derecelerinin belirlenmesi için kullanılabileceğini açıkta belirtmektedir. Başvuru metninin dilde ölçme değerlendirme alanında oynayabileceği rolün önemi yadsınamazken, başvuru metnine göre hazırlanan yeni sınavlar ve bünyesinde düzenlemeler yapılan hâlihazırdaki sınavlara eleştirel bir gözle bakmak gereklidir. Bu açıdan, mevcut çalışma, dilde ölçme ve değerlendirme alanına yönelik olarak başvuru metninin sahip olduğu potansiyel problemleri ve uygulamaya dönük unsurları incelemeyi amaçlamakta, dilde ölçme alanında çalışan paydaşlar ve dil öğretmenleri için uygulamaya dönük çıkarımları tartışmaktadır.

Anahtar Sözcükler: Avrupa Dilleri Öğretimi Ortak Çerçeve Programı; yabancı dil öğretiminde ölçme ve değerlendirme; test geliştirme; uyarlama

AUTHOR BIODATA

Paşa Tevfik Cephe is an associate professor of English Language Teaching at Gazi University, Gazi Faculty of Education in Ankara, Turkey, where he teaches special teaching methods, curriculum development and syllabus design, and research methodology courses. Dr. Cephe has published many articles and book chapters on several issues in ELT, besides carrying out projects affiliated with the Ministry of Education.

Tuğba Elif Toprak is a PhD candidate and research assistant in the Department of Foreign Language Education, Gazi University in Ankara, Turkey. Her research interests include language testing, specifically diagnostic testing, cognitive psychology, and psycholinguistics.