From Common European Framework to Classroom Application: the English Profile solution

Anthony Green

Centre for Research in English Language Learning and Assessment, University of Bedfordshire, UK

tony.green@beds.ac.uk

Abstract

The English Profile is a multidisciplinary programme of research directed towards answering the Council of Europe's (2005) call for a set of Reference Level Descriptions (RLDs) for English linked to the Common European Framework of Reference for Languages (CEFR). The Council of Europe has issued guidelines setting out broad parameters for RLD development. This paper discusses how RLD can be developed for English in relation to the aims of the CEFR, incorporating consideration of critical voices, reports on the experiences of users of the CEFR and a review of currently operational RLDs for English: the *Threshold* series. On the basis of these sources, recommendations are made for the ongoing development of the English Profile. Specifically, a process is suggested for working with the CEFR to generate locally apposite educational tools..

Keywords

Common European Framework of Reference; CEFR; educational objectives; assessment; curriculum development; proficiency

Introduction

The Common European Framework of Reference for Languages (CEFR) (Council of Europe [CoE], 2001) offers 'a common basis for the elaboration of language syllabuses, curriculum guidelines, examinations, textbooks etc.' (CoE, 2001: 1). It is clear that the appeal of the framework now extends beyond Europe and users from around the world feel the need to engage with the framework either because it can offer them guidance or because they are compelled to by external agencies, or both.

Despite criticism by Fulcher 2004; Krumm 2007; and Jones & Saville 2009) that some stakeholders use the framework as a means of imposing harmonisation – applying it to language programmes 'as a hammer gets applied to a nail' in Jones and Saville's (2009: 54) phrase – this is not its intended purpose. Rather the framework is intended as a heuristic tool that will allow for and capture 'the possible diversity of learning aims and the variety to be found in the provision of teaching' (CoE 2001: 138). It is emphasised throughout the CEFR that it is a resource for consultation that 'can be applied, with such adaptations as prove necessary, to particular situations' (CoE 2001: 7).

The intended capacity of the framework both to describe multiple levels of ability and to cater to multiple contexts for language use is expressed by Richterich & Schneider (1992) through the concepts of *horizontality* and *verticality*. Horizontality is explained in the following terms: '[the] description and clarification of multi-dimensional content in terms of linguistic, social and cultural attainment, communication situations, or partial skills such as reading comprehension of texts of a certain type' (Richterich & Schneider 1992: 44). Building on the 'descriptive scheme' of the CEFR, Can Do descriptors can be grouped horizontally in ways that are meaningful for specific audiences – learners, teachers, employers, government agencies – for use in diverse contexts. Learners placed at the same global level, but with different needs, may set themselves objectives based on – or be assessed against – descriptors drawn from entirely separate CEFR tables.

The range of categories is said to allow for the construction of scales that reflect specific contexts for language use within the four specified 'domains': 'personal', 'public', 'occupational' and 'educational' (CoE 2001: 45) and in relation to different areas of the user's competence.

On the vertical dimension, descriptors are grouped into 'performance level descriptions'. A global scale, suggests one possible holistic summary of these levels and is presented in Table 4 of the CEFR in a form that might convey broad information to the 'non-specialist user' (CoE 2001: 24). More detailed alternatives are suggested for learner self-assessment (CoE 2001: 26–27) and for the assessment of spoken performance (CoE 2001: 28–29). The self assessment scale presents distinctions between the skills of reading, listening,

writing, spoken interaction and spoken production and the spoken performance scale further distinguishes within spoken language between such 'qualitative aspects of language use' as 'accuracy', 'fluency' and 'coherence' (CoE 2001: 25).

The flexibility and local adaptability of the scheme is stressed throughout. In the 'branching approach' suggested, a broader distinction can be made between three superordinate levels of learner (A: basic, B: independent and C: proficient). Finer distinctions can be made within the CEFR levels so that relatively small gains in language proficiency can be captured and reported.

The framework is said to be capable both of informing mastery decisions and of locating a performance on a continuum of proficiency (CoE 2001: 184). In other words, the descriptors are intended to be used both as a basis for specifying tasks that a learner at that CEFR level might be expected to work towards or succeed in performing – a 'constructor-oriented' purpose in Alderson's (1991) terms – and in providing differentiated descriptions of the quality of linguistic performance that can be used as rating scales to assign learners (all of whom might perform the same task) to the most appropriate level – an 'assessor-oriented' purpose (Alderson 1991).

1 Reference Level Descriptions and the CEFR

The Council of Europe (2005) has called for the development of Reference Level Descriptions (RLDs). This is an acknowledgement of the need to interpret and elaborate the necessarily broad descriptions of levels in the CEFR so that they may more readily inform practices relating both to specific languages and to specific applications (such as syllabus design or test development) (CoE 2005).

In common with other elements of what has come to be known as the CEFR toolkit, RLDs are intended to assist users in employing the CEFR to meet their local needs as language learners and educators.

The guidelines specify that RLDs should provide 'inventories of the linguistic realisations of general notions, acts of discourse and specific notions/lexical elements and morpho-syntactic elements' (CoE 2005: 5) that ground the CEFR descriptors. There is no requirement to limit the descriptors employed to those used in the CEFR; indeed, developers are encouraged to incorporate descriptors from European Language Portfolio (ELP) models. Each RLD should explain the process by which these inventories are arrived at, the knowledge of a linguistic form expected of learners at a given level (receptive or productive) and the relationships between the lists presented.

RLDs are thus intended to mediate between the CEFR and specific contexts for its use. They offer meaningful illustrative learning objectives that more fully operationalise the CEFR descriptions for users by providing linguistic exponents. In considering how best to proceed in developing RLDs, our first step should be to consider how learning objectives are presented in the CEFR and the extent to which these objectives already meet the needs of CEFR stakeholders as revealed through the available literature.

2 The action-oriented approach

The action-oriented approach of the CEFR prioritizes what learners are able to accomplish when using a language over their knowledge about language. Learning objectives are conceived in terms of language activities and tasks – 'any purposeful actions considered by an individual as necessary in order to achieve a given result in the context of a problem to be solved, an obligation to fulfil or an objective to be achieved' (CoE 2001: 10).

To assist in the process of setting objectives for language learning, teaching and assessment, the CEFR presents the user with conceptual questions for consideration and a bank of 'illustrative' descriptors arranged according to the descriptive scheme into 54 scales, each relating either to an aspect of a 'competence' (e.g. scales for 'vocabulary range' and 'vocabulary control' [CoE 2001: 112] are aspects of 'lexical competence' [CoE 2001: 30]) or to a language 'activity' (e.g. for oral production, 'Sustained monologue: describing experience', 'Sustained monologue: putting a case (e.g. in debate)', 'Public announcements', 'Addressing audiences').

The illustrative descriptors are arranged into six levels of proficiency. According to the CEFR, 'The intention of providing a concrete illustrative set of descriptors, together with criteria and methodologies for the further development of descriptors, is to help decision-makers design applications to suit their contexts' (CoE 2001: 36).

Reflecting the action oriented approach, the descriptors provided take the form of positively worded Can Do statements describing activities that the learner might carry out in the target language. They are said to draw on three main sources: '(a) the theoretical work of the [CEFR] authoring group, (b) the analysis of existing scales of proficiency and (c) the practical workshops with teachers' (CoE 2001: 30). It is claimed that the scaling of these descriptors based on the rating of performance samples using the Rasch model (Bond and Fox,

2007) creates an empirically derived scale of language ability.

3 Interpreting Can Do descriptors

The Can Do descriptors used in developing the illustrative scales and determining the cut points between performance level descriptions (North 2000). were adapted from the ACTFL Proficiency Guidelines and related proficiency scales such as the Australian Second Language Proficiency Ratings (Ingram 1990) together with educational objectives-based schemes such as the English National Curriculum: Modern Languages (1991) and the Eurocentres Scale of Language Proficiency (1993). 30 schemes altogether were included giving an initial pool of 1,679 descriptors in total (North 2000).

Because the illustrative descriptors constitute independent, criterion statements which have been calibrated to the levels concerned, they can be used as a source to produce both a checklist for a particular level, as in some versions of the Language Portfolio, and rating scales or grids covering all relevant levels, as presented in Chapter 3 of the CEFR, for self-assessment in Table 2 and for examiner assessment in Table 3 (CoE 2001: 189).

The CEFR approach was innovative and attracted widespread praise for its rigor, but it has also drawn criticism. Hudson (2005) questions the claim of empiricism, observing that 'whereas the descriptors were empirically scaled based on performance ratings, the particular descriptors were not subsequently cast as actual test prompts and then calibrated again to determine if they still scale hierarchically' (p.218). North (2000) acknowledges the further objection that the scales are empirical only to the extent that they calibrate teacher perceptions: they are not empirically derived from L2 learner data (Hulstijn 2007). It is clear that reference level descriptions will need to address the links between teacher perceptions as operationalised in the scales and observable learner performance.

The criticisms made of the ACTFL Proficiency Guidelines (Lantolf & Frawley 1985; Savignon 1985; Kramsch 1986) that they were not based on research into the nature of second language acquisition have been repeated with respect to the CEFR by Hulstijn (2007) among others. Although, unlike the ACTFL Proficiency Guidelines, the CEFR is based on a componential model of communicative language ability, Hulstijn (2007: 666) notes the continuing lack of 'empirical evidence that, in following the overall oral proficiency scale, all learners first attain the functional level of A1, then the level of A2, and so on, until they reach their individual plateau'.

The tasks that learners are asked to perform have a substantial impact on the nature and quality of the language that they will be able to produce.

Further work is needed to tease out the relationship between elements of linguistic quantity and quality. The complex relationship between the two may partly explain the feeling that the higher levels of the CEFR, which take account of the academic and professional orientation of much language learning at advanced levels, are inappropriate or unattainable for certain groups of learners – especially young learners. Activities envisaged at these levels 'lie beyond the cognitive and experiential range of children and the great majority of adolescents' (Little 2007: 651).

A further issue raised by Green (2011) is that although a list of external conditions is provided (CoE 2001: 46–47), these are not incorporated into or explicitly related to the Can Do descriptors that define the levels. Weir (2005) has pointed out that this must seriously undermine the interpretability of the levels as it is unclear which conditions should apply when we judge what learners 'can do'.

The approach taken by the CEFR is contrasted by McNamara & Roever (2006) with schemes such as the Australian Certificates in Spoken and Written English (CSWE) or the related Canadian Language Benchmarks (CLB 2000) which provide elaboration and sample tasks to exemplify levels. Such elaboration is intended to help test developers, teachers responsible for assessment and learners themselves to arrive at comparable understandings of learner abilities. The CEFR descriptions, in contrast, are used in a much wider range of settings, but are left far more open to interpretation.

4 Adopting and adapting the CEFR

The lack of guidance on criteria and conditions in the CEFR leads directly to questions about comparability. In posing the question, 'how does one know for certain that a test of Greek calibrated at level B1 in Finland is equivalent to a test of Polish considered to be at level B1 in Portugal?', Bonnet (2007: 670) points to two threats to comparability: local norms and inter-linguistic variation. In the absence of elaboration, exemplification and, most crucially, moderation of standards, it is likely that users in one setting may interpret the illustrative descriptors differently to users in another.

It is possible, even likely, that different interpretations of levels will develop so that, to give an entirely fictional example, a learner judged to be at level B1 in one school in Finland might be rated as level C1 at another in Portugal. The risks associated with such inconsistency are well

illustrated by Crossey's (2009) account of the NATO STANAG 6001 scheme intended to provide agreed international language standards for military personnel. Faith in the scheme was undermined when it became clear that learners certified as being at a given level in one context did not satisfy the criteria as interpreted in another.

It is not surprising that users feel the need for more detailed specifications of level. Martyniuk & Noijons (2007) in their survey of the use of the CEFR across Europe found that users 'stress the need for general clarification (such as comments on theoretical concepts, examples and good illustrations, sets of tasks for use in specific contexts'.

The further development of sample materials provided by the Council of Europe or of open-access schemes such as DIALANG will no doubt help to ground the CEFR levels in a way that meets these needs. But this is not without risk: the more detail that is given, the greater the risk that *illustrative* tasks become *required* tasks and this distinction must be made clear.

To guard against such risks, it is important that RLDs should be pluralistic. They should incorporate a mechanism for ongoing interaction between users so that common understandings can be fostered and maintained by collective discussion, not imposed by a central agency.

5 Operationalizing the CEFR for local purposes

Green (2011) shows how locally tailored descriptors can be generated, calibrated and integrated into the CEFR descriptive scheme and reference levels by extending the original CEFR methodology.

Green (2011) proposes that Can Do descriptors should include, or be supported by, specifications of operative conditions and performance standards. In an attempt to supplement the limited number of descriptors available for the C1 and C2 levels of the CEFR, he synthesised descriptions of high level objectives from educational materials in global use, mapped these to a Can Do template and used a survey of educators to calibrate them to the CEFR levels.

However, he argues that better meeting the needs and expectations of users will require more than the generation of Can Do statements. A number of informational layers will be needed to elaborate Can Do descriptors for use in curricula or assessment schemes where shared interpretations are important. These layers could be made up of components such as the following:

1. Generative Can Do frames: these set out which of the elements of the CEFR descriptive scheme may interact in shaping the difficulty of defined language activities and tasks.
2. Grids of criterial features: lists of features that impact on the difficulty of the relevant language activities and estimates of how these might affect level estimates. How is performance of a task at one level most clearly differentiated from performance of the same or similar tasks at an adjacent level?
3. Glosses: definitions and elaborations of key words used in the reference level descriptions.
4. Commentaries: discussions of how the components of the reference level descriptions might be interpreted in relation to specified tasks.
5. Sample tasks: examples of (receptive and productive) tasks that learners at different levels might be expected to carry out, with commentary explaining how these relate to the interpretation of the level descriptions.
6. Sample performances: examples of learner performance – recordings or scripts – illustrating the interpretation of the level descriptions.

Where material of this kind is generated, it can serve both to explain the relatively abstract descriptions to teachers, learners and others who need to interpret outcomes, but it can also improve the consistency with which descriptors are interpreted.

It is becoming clear from the experience with the CEFR that it might be damaging to view it as a static and universal delineation of English language education. It must be both flexible and open to challenge: it should encourage debate and reflection and allow for regular revision.

In accordance with the approach outlined in the CEFR, the relative difficulty of the multiple means of realising language activities must be considered both from a social perspective and from a cognitive perspective allowing for the diversity of contexts in which organised language learning occurs.

Any given activity involving language use requires the activation of cognitive language processes, with the 'many-to-many' (Hawkins & Filipovič 2011) relationship between linguistic form and function opening a wide variety of choices to the user. Processing occurs 'in relation to themes, in specific domains' (Council of Europe 2001: 9). This social context of themes and domains places constraints on the range of choices open to the user and on the way in which utterances will be understood. Only by addressing the interaction between the cognitive and the social does it become

possible to adequately triangulate the relative demands of language learning tasks for the intended purposes of objective-setting, teaching and testing.

References

Alderson, J. C. (1991). Bands and scores. In Alderson, J. C. & North, B. (eds.) *Language testing in the 1990s*. London: Macmillan, 71–86.

Bonnet, G. (2007). The CEFR and education policies in Europe. *Modern Language Journal* 91(4): 669–672.

Bond, T. G. & Fox, C. M. (2007). *Applying the Rasch Model: Fundamental measurement in the human sciences*. Mahwah, NJ: Lawrence Erlbaum.

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

Council of Europe (2005). *Draft Guide for the Production of RLD: Version 2*. Strasbourg: Language Policy Division.

Crossey, M. (2009). The role of micropolitics in multinational, high-stakes language assessment systems. In Alderson, J. C. (ed.), *The politics of language education: Individuals and institutions*. Bristol: Multilingual Matters, 147–164.

Fulcher, G. (2004). Deluded by artifices? The Common European Framework and harmonization. *Language Assessment Quarterly* 1(4), 253–266.

Green, A. (2011). *Language Functions Revisited: Theoretical and empirical bases for language construct definition across the ability range.* Cambridge: Cambridge University Press.

Hawkins, J.A. & Filipovič, L. (2011). *Criterial Features in L2 English.* Cambridge: Cambridge University Press.

Hudson, T. (2005). Trends in assessment scales and criterion-referenced language assessment. *Annual Review of Applied Linguistics* 25, 205–227.

Hulstijn, J. H. (2007). The shaky ground beneath the CEFR: Quantitative and qualitative dimensions of language proficiency. *Modern Language Journal* 91(4), 663–667.

Jones, N. & Saville, N. (2009). European language policy: Assessment, learning and the CEFR. *Annual Review of Applied Linguistics*, 29, pp. 51–63.

Kramsch, C. (1986). From language proficiency to interactional competence. *Modern Language Journal* 70,(4) 366–72.

Krumm, H-J. (2007). Profiles instead of levels: The CEFR and its (ab)uses in the context of migration. *Modern Language Journal* 91(4), 667–669.

Lantolf, J. & Frawley, W. (1985). Oral proficiency testing: A critical analysis. *Modern Language Journal* 69(4), 337–345.

Little, D. (2007). The Common European Framework of Reference for Languages: Perspectives on the making of supranational language education policy. *Modern Language Journal* 91(4), 645–655.

Martyniuk, W and Noijons, J (2007). *Executive summary of results of a survey on the use of the CEFR at national level in the Council of. Europe Member States*. Strasbourg: Council of Europe.

McNamara, T. F. & Roever, C. (2006). *Language testing: The social dimension*. Oxford: Blackwell.

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

Richterich, R. & Schneider, G. (1992). Transparency and coherence: why and for whom? In North (Ed.), 43–50.

Savignon, S. (1985). Evaluation of communicative competence: The ACTFL Provisional Proficiency Guidelines. *Modern Language Journal* 69(2), 129–134.

Weir, C. J. (2005). Limitations of the Council of Europe's Framework of reference (CEFR) in developing comparable examinations and tests. *Language Testing* 22(3), 281–300.