ResearchNotes
UNIVERSITY of CAMBRIDGE ESOL Examinations
English for Speakers of Other Languages
EXTRACT FROM RESEARCH NOTES ISSUE 24 / MAY 2006 / PAGES 2–5 24/01

Cambridge ESOL exams and the Common European Framework of Reference (CEFR)

LYNDA TAYLOR AND NEIL JONES, RESEARCH AND VALIDATION GROUP

Introduction

A previous *Research Notes* article explored issues of test comparability and the role of comparative frameworks as communicative tools (Taylor 2004). One framework which has a growing role for language testers is the Common European Framework of Reference (CEFR; Council of Europe 2001). At Cambridge we are often asked about the relationship between our ESOL exams and the CEFR; the nature of this relationship can be considered from four complementary, sometimes overlapping, perspectives.[1]

The historical perspective

The origins of the CEFR date back to the early 1970s when the Council of Europe sponsored work within its Modern Languages Project to develop the Waystage and Threshold levels as sets of specified learning objectives for language teaching purposes. These two levels were designed to reflect achievable and meaningful levels of language competence, at a relatively low proficiency level, and to form part of a European unit/credit system for adult language learning. They defined levels of functional competence among language users forming the basis for curriculum, syllabus, and later assessment design.

In the late 1980s Cambridge was one of several stakeholder organisations (with the British Council and BBC English) to provide funding and professional support for revising Threshold and Waystage (Van Ek and Trim 1998a, 1998b); the revised level descriptions underpinned test specifications for a revised PET exam in the mid 1980s and a new KET exam in the early 1990s.

Linguistic and functional description of a third, higher proficiency level began in the 1990s, with support and participation on this occasion from the Association of Language Testers in Europe (ALTE); work on this third level took account of FCE and led to the publication of Vantage in 1999 (Van Ek and Trim 2001). As work extended on level descriptions for English, so the concept of a framework of reference levels began to emerge and to take on a more concrete form.

The conceptual perspective

In part, emergence of a framework formalised conceptual levels with which ELT learners, teachers and publishers had operated for some years – with familiar labels such as 'intermediate' or 'advanced'. Dr Brian North, one of the CEFR's authors, confirms its origins in traditional English Language Teaching levels:

> *The CEFR levels did not suddenly appear from nowhere. They have emerged in a gradual, collective recognition of what the late Peter Hargreaves (Cambridge ESOL) described during the 1991 Rüschlikon Symposium as "natural levels" in the sense of useful curriculum and examination levels.*
>
> *The process of defining these levels started in 1913 with the Cambridge Proficiency exam (CPE) that defines a practical mastery of the language as a non-native speaker. This level has become C2. Just before the last war, Cambridge introduced the First Certificate (FCE) – still widely seen as the first level of proficiency of interest for office work, now associated with B2. In the 1970s the Council of Europe defined a lower level called "The Threshold Level" (now B1), originally to specify what kind of language an immigrant or visitor needed to operate effectively in society. Threshold was quickly followed by "Waystage" (now A2), a staging point half way to Threshold. The first time all these concepts were described as a possible set of "Council of Europe levels" was in a presentation by David Wilkins (author of "The Functional Approach") at the 1977 Ludwighaven Symposium…(North 2006:8).*

Cambridge's upper-intermediate level CAE exam, introduced in 1991, helped bridge the gap between FCE and CPE and was proposed as C1. Lastly, a lower Breakthrough level was proposed as A1. These six levels (A1-C2) thus constituted a 'language ladder', providing a pathway for upward progression in language teaching and learning with explicit opportunities to evaluate and accredit learning outcomes along the way. The Cambridge Main Suite exams (KET, PET, FCE, CAE and CPE) were already providing well-established and recognised accreditation 'stepping stones' along this pathway.

Emergence of these common reference levels, with their contributory elements such as language courses, public examinations, and published coursebooks, was formally confirmed through the Common European Framework project; managed between 1993 and 1996 by the Council of Europe with significant input from the Eurocentres organisation, the overarching aim was

1. This article is based on a presentation given at IATEFL Harrogate in April 2006 and we are grateful to Dr Brian North for his helpful comments on an early draft.

to construct a common framework in the European context which would be transparent and coherent, to assist a variety of users in defining language learning, teaching and assessment objectives. A major strength was that it would build upon the shared understanding which already existed among teachers and other ELT stakeholders in the European context, but would also resolve some difficulties of relating language courses and assessments to one another; it would provide a common meta-language to talk about learning objectives and language levels and encourage practitioners to reflect on and share their practice. It's worth remembering that this took place in a larger context where notions of a socio-political and economic community in Europe were rapidly taking shape; an early motivation for revising Waystage and Threshold in the late 1980s had been their relevance to educational programmes of language learning for European citizenship.

Notions of framework development linked to language learning progression were nothing new. Wilkins' 1977 set of levels has already been referred to. In the UK context, the English Speaking Union (ESU) set up its 'framework project' in 1985 to devise a comprehensive frame of description for comparing the various examinations of the main English language boards (Taylor 2004). In the wider context of Europe, ALTE members were also by the early 1990s working systematically to co-locate their qualifications across different European languages and proficiency levels within a shared framework of reference. The aim was to develop a framework to establish common levels of proficiency in order to promote the transnational recognition of certification in Europe. The process of placing ALTE members' exams on the framework was based on content analysis of the tests, the creation of guidelines for the quality production of exams, and the development of empirically validated performance indicators or Can Do statements in different European languages (see ALTE website www.alte.org). The resulting five-level ALTE Framework developed simultaneously during the mid-1990s alongside the six-level CEFR published in 1997. Since the two frameworks shared a common conceptual origin, similar aims – transparency and coherence – and comparable scales of empirically developed descriptors, Cambridge ESOL and its ALTE partners decided to conduct several studies to verify their alignment. This was achieved mainly through the ALTE Can Do Project in 1998-2000 (see below). Following publication of the CEFR in 2001 the ALTE members adopted the six CEFR levels (A1-C2).

One of the strengths of this conceptual approach to framework development has undoubtedly been its 'organic' development. Even in 1991, qualifications existed for other languages that could also be confidently associated with what were to become the CEFR and ALTE levels, including: the new advanced level DALF (Diplôme Approfondi de Language Française) at C1; the Zertifikat Deutsch (ZD) at Threshold (B1); and the Kleines Deutsches Sprachdiplom (KDS) commonly considered an equivalent to Cambridge's CPE (C2).

The empirical perspective

Shared understanding among teachers, publishers and language testers enabled the framework concept to function quite well without extensive underpinning from measurement theory and statistics; but measurement theory has become increasingly important as attempts have been made to validate aspects of the CEFR empirically (North and Schneider 1998, North 2000a) and to link assessments to it (North 2006b).

Syllabus designers, coursebook publishers and language test providers worldwide, including Cambridge ESOL, seek to align their exams to the CEFR for reasons of transparency and coherence; claims of alignment can also assist in marketing communications to try and gain a competitive edge. However, any claim of alignment needs to be examined carefully; simply to assert that a test is aligned with a particular CEFR level does not necessarily make it so, even if that assertion is based on an intuitive or reasoned subjective judgement. To some extent, alignment can be achieved historically and conceptually as we have seen, but empirical alignment requires more rigorous analytical approaches; appropriate evidence needs to be accumulated and evaluated.

The ALTE Can Do Project (Jones 2001, 2002) was one of the empirical approaches used by Cambridge ESOL for aligning its original five levels with the six-level CEFR. Other empirical support for alignment comes from Cambridge's item-banking methodology underpinning our approach to all test development and validation (Weir and Milanovic 2003). The Cambridge-TOEFL Comparability Study, conducted in 1987-90 (Bachman et al 1995) highlighted how far the UK-based assessment tradition had relatively underplayed the psychometric dimension; for Cambridge ESOL this established an empirical imperative and we invested heavily in approaches and systems to address measurement issues such as test reliability and version comparability. Latent trait methods have been used since the early 1990s to link the various Cambridge levels onto a common measurement scale using a range of quantitative approaches, e.g. IRT Rasch-based methodology, alongside qualitative research methods.

More recently, Cambridge ESOL has supported the authoring and piloting of the Council of Europe's Manual Relating Language Examinations to the CEFR (Figueras et al 2005) which presents a linking process based on three sets of procedures:

Specification of the content and purpose of an examination

Similar procedures were conducted when the PET and KET test specifications were originally based upon Threshold and Waystage levels, and the ALTE partners' exams were aligned within the ALTE Framework; an extensive range of documentation for all our exams (test specifications, item writer guidelines, examiner training materials, test handbooks and examination reports) assists in specifying the content and purpose of existing and new exams with direct reference to the CEFR.

Standardisation of interpretation of CEFR levels

Suitable standardised materials are needed for assessment personnel and others to benchmark their tests against CEFR levels. Cambridge has helped develop such materials by supplying calibrated test items and tasks from our Main Suite Reading and Listening test item banks together with exemplar Speaking and Writing test performances from our writing examiner coordination packs and Oral Examiner standardisation materials at each CEFR

level; a set of benchmarking materials, incorporating both classroom-based and test-based materials, is now available from the Council of Europe on CD or DVD.

Empirical validation studies

Empirical validation studies are a greater challenge sometimes requiring specialist expertise and resources; Cambridge ESOL is among a relatively small number of examination providers undertaking this sort of research, partly through our routine item-banking and test calibration methodology and also through instrumental research and case studies such as the Common Scale for Writing Project (Hawkey and Barker 2004).

The evolutionary perspective

The CEFR remains 'work in progress'; it will continue to evolve as experience grows among those who use it in various ways and contexts, and as they reflect on that use. For many it already provides a useful frame of reference, offering practical guidance for their thinking and doing. Others have expressed concerns about its application: within the language testing community some fear use of the CEFR as an instrument for 'harmonisation' of policy/practice (Fulcher 2004); others question how far the CEFR provides a suitable instrument for operational test development (Weir 2005). In response, the CEFR authors emphasise the original intention of the Framework as a means of valuing and encouraging diversity, and remind us that the CEFR is not a 'cookbook' or 'how to' document. Perhaps the real value of the CEFR lies in it being used as a heuristic rather than prescriptively; it needs to be interpreted thoughtfully and intelligently if it is to be meaningful and have local validity.

Another useful role for the Framework in assessment could be in matters of quality assurance, not just to improve systems and procedures but to support the growing professionalisation of personnel and institutions involved in language learning, teaching and assessment. North (2006) notes that the scheme outlined in the Manual 'reflects the three step process of any Quality Management System (Design, Implementation, Evaluation)'. This view echoes Cambridge ESOL's long-standing commitment to addressing quality assurance issues. In the early 1990s ALTE produced its professional Code of Practice and has since then elaborated the concept of quality assurance in language testing by developing quality management instruments. Like the CEFR, the ALTE Code of Practice offers the practitioner community a common frame of reference and a shared meta-language for reflecting on and evaluating policy and practice – ensuring the door is always open for improvement.

Since 2001, the CEFR has also been a source of inspiration or a catalyst for other initiatives; one is the innovative European Language Portfolio (ELP) developed to support the language learning and teaching community with input from the EAQUALS organisation and the ALTE partners; another is the recently launched English Profile Project to develop a comprehensive set of Reference Level Descriptions for English using the CEFR levels as a springboard.

Conclusion

Today the CEFR plays a key role in language and education policy within Europe and the wider world – perhaps in ways not originally envisaged by its authors. Within Europe it is believed to serve policy goals of fostering linguistic diversity, transparency of qualifications, mobility of labour, and lifelong language learning. Beyond Europe it is being adopted to help define language proficiency levels with resulting implications for local pedagogy and assessment. For Cambridge ESOL it offers a valuable frame of reference for our work and for our stakeholder community; as it evolves, we look forward to continuing to make an appropriate professional contribution to its development.

Could it be argued that Cambridge ESOL exams 'embody' the Common European Framework? That will be for others to judge based on evidence presented here and elsewhere. It partly depends on how the word 'embody' is defined; but there does exist a growing body of evidence to support a claim that Cambridge exams contain and express the CEFR as an important feature, that they include the CEFR as part of their structure, and that they express or represent the CEFR in a variety of ways. Such embodiment is a natural outcome of several factors, such as historical legacy, conceptual synergy, and empirical underpinning. Extending the biological metaphor, we could envisage how the relationship between the CEFR and Cambridge ESOL exams will continue to evolve, partly due to the genetic makeup of the relationship itself and also as a result of external environmental factors in a changing world.

To celebrate his 80th birthday in 2004, Professor John Trim, one of the authors of the CEFR, was interviewed for Language Assessment Quarterly. In the interview, he describes the aspirations behind the Framework: 'What we were aiming at was something which will be a common reference point that people working in different fields and people using it for entirely different things and in very different ways could refer to in order to feel that they were part of a common universe' (Saville 2005:281). This focus on individual practitioners as the agents of activity is a welcome reminder that it is people, rather than frameworks, systems, or procedures, who are – or who should be – at the heart of what happens in language learning, teaching and assessment, i.e. learners, teachers, teacher trainers, course and syllabus designers, textbook writers, language test providers – anyone who is a stakeholder in the ELT or ESOL constituency, or who is a member of another language learning community.

Ultimately, it may be unhelpful to talk about 'embodiment' in relation to a course syllabus or an assessment tool; of greater interest and importance, both to the developers of the CEFR and to Cambridge ESOL, are surely the populations of human beings directly involved in language learning, teaching and test-taking, whether at the group or the individual level. The quality of the relationship between the CEFR and Cambridge ESOL exams is perhaps best judged by the extent to which together they enable language learning to flourish, encourage achievements to be recognised and so enrich the lives of individuals and communities.

References and further reading

Bachman, L F, Davidson, F, Ryan, K and Choi, I C (1995) *An investigation into the comparability of two tests of English as a foreign language: The Cambridge-TOEFL comparability study*, Cambridge: UCLES/Cambridge University Press.

Council of Europe (2001) *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*, Cambridge: Cambridge University Press.

Figueras, N, North, B, Takala, S, Verhelst, N and Van Avermaet, P (2005) Relating Examinations to the Common European Framework: a Manual, *Language Testing* 22 (3), 1–19.

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

Hawkey, R & Barker, F (2004) Developing a common scale for the assessment of writing, *Assessing Writing* 9/2, 122–159.

Jones, N (2001) The ALTE Can Do Project and the role of measurement in constructing a proficiency framework, *Research Notes* 5, 5–8.

—(2002) Relating the ALTE Framework to the Common European Framework of Reference, in Alderson, J C (Ed.), *Common European Framework of Reference for Languages: Learning, Teaching, Assessment. Case Studies*, Strasbourg: Council of Europe Publishing, 167–183.

North, B (1995) The development of a common framework scale of descriptors of language proficiency based on a theory of measurement, *System* 23, 445–465.

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

—(2000b) Linking Language Assessments: an example in a low-stakes context, *System* 28, 555–577.

—(2006) The Common European Framework of Reference: Development, Theoretical and Practical Issues, paper presented at the symposium 'A New Direction in Foreign Language Education: The Potential of the Common European Framework of Reference for Languages', Osaka University of Foreign Studies, Japan, March 2006.

North, B and Schneider, G (1998) Scaling Descriptors for Language Proficiency Scales, *Language Testing* 15 (2), 217–262.

Saville, N (2005) An interview with John Trim at 80, *Language Assessment Quarterly* 2 (4), 263–288.

Taylor, L (2004) Issues of test comparability, *Research Notes* 15, 2–5.

Van Ek, J A and Trim, J L M (1998a) *Threshold 1990*, Cambridge: Cambridge University Press.

—(1998b) *Waystage 1990*, Cambridge: Cambridge University Press.

—(2001) *Vantage*, Cambridge: Cambridge University Press.

Weir, C J (2005) Limitations of the Common European Framework for developing comparable examinations and tests, *Language Testing* 22 (3), 281–300.

Weir, C J and Milanovic, M (2003) *Continuity and Innovation: Revising the Cambridge Proficiency in English examination 1913–2002*, Studies in Language Testing Vol. 15, Cambridge: UCLES/Cambridge University Press.

This is an extract from *Research Notes*, published quarterly by University of Cambridge ESOL Examinations.
Previous issues are available on-line from www.CambridgeESOL.org.