The English Profile Programme – the first three years

Nick Saville Director, Research and Validation, University of Cambridge ESOL Examinations

Roger Hawkey Consultant, University of Cambridge ESOL Examinations

Abstract

The English Profile Programme is a major inter-disciplinary, research-based collaborative project to develop detailed reference level descriptions (RLDs) of the English of L2 language learners, linked to the general principles and approaches of the Council of Europe. The Common European Framework of Reference (CEFR) provides language proficiency bands against which the profile is categorised. Large samples of writing and speech are required for the Programme, with the Cambridge Learner Corpus providing data for the initial phase.

Keywords: Reference Level Descriptions, corpora, socio-cognitive approach, English Profile Programme

1. Introduction

This paper will provide an overview of English Profile, a collaborative programme of interdisciplinary research set up to produce reference level descriptions (RLDs) for English linked to the general principles and approaches of the Council of Europe. The intended output of English Profile is a 'profile' of the English language levels of learners in terms of the six language proficiency bands of the Common European Framework of Reference (CEFR). These common reference levels subdivide the conventional three language level categories – basic, intermediate and advanced - into six, A1 (Breakthrough) and A2 (Waystage); B1 (Threshold) and B2 (Vantage); C1 (Effective Operational Proficiency) and C2 (Mastery). However, as Milanovic (2009: 5) reminds us, the indicative levels of the CEFR were designed to be applied to *all* the diverse languages of Europe rather than one particular language:

> The CEFR is neutral with respect to language and, as the <u>common</u> framework, must by necessity be underspecified for all languages. This means that specialists in the teaching or assessment of a given language . . . need to determine the linguistic features which increasing proficiency in the language entails . . . Such features are peculiar to each language and so the CEFR must be adapted to accommodate the language in question.

Referring to the English Profile itself, Milanovic continues:

> A major objective of English Profile is to analyse learner language to throw more light on what learners of English can and can't do at different CEFR levels, and to assess how well they perform using the linguistic exponents of the language at their disposal (i.e. using the grammar and lexis of English).

Indeed. The emerging English Profile reference level descriptions are designed specifically for the *English* language (see further below).

The Council of Europe descriptions at ascending levels already form the basis for specifications of course design and levels of competence both in national and international contexts as well as for the design of multimedia courses. It is significant that these level descriptions tend to emphasise, as key features in their descriptions of the language typifying a CEFR level, elements such as grammar, language notions and functions, the lexicon, and discourse structure. The influence of these Council of Europe analytic features on the English Profile model is clearly shown below and in other articles in this journal.

2. The CEFR and its Reference Level Descriptions (RLDs)

In this section, we provide a brief historical overview of the work of the Council of Europe on which the English Profile Programme is based.

In 2001, the Council of Europe published the finalised version of its Common European Framework, which had been in preparation for ten years since the Rüschlikon Conference of 1991. The CEFR was the result of much consultation on several draft versions (1996, 1998), which were widely circulated before the final text was ready.

However, the story goes back much further, to the early 1970s. The CEFR authoring team built their approach on principles which had seen the emergence of the *communicative approach to language teaching* in Europe, and to the publication by the Council of Europe of ground breaking texts such as the *Threshold* and *Waystage* specifications of language levels (van Ek, 1975a, 1975b; van Ek & Trim, 1990a, 1990b, 2001). These titles are sometimes referred to collectively as the T-series.

Since its publication, the CEFR has certainly achieved far more than the authors might have imagined; it has been disseminated very widely around the globe, translated into 37 languages and is now a common reference instrument for organising language teaching and assessment in many countries (see Little, 2006 for an overview). The fact that each level is described within a coherent framework of six reference levels has been one of the most influential aspects of the CEFR. The framework has generated a great deal of discussion in the fields of curriculum development, language teaching and assessment, though perhaps with an over emphasis on the levels rather than the rest of the 2001 text (see Coste, 2007).

In their original Council of Europe context the descriptions of the levels through *illustrative scale descriptors* needed to be general so that they could be applied to the learning and teaching of *any* language. It was always recognised, as Milanovic suggests above, that the descriptors would need to be supplemented by *language-specific* illustrations and guidance on how to apply the framework to different languages in local contexts.

So as a central part of its work since 2001, the Council of Europe has encouraged the production of instruments to complement the CEFR, including RLDs for national and regional languages. Such RLDs represented a new generation of descriptions which would identify the specific forms of any given language (words, grammar, etc.), the mastery of which would correspond to the competences defined by the CEFR at each of the six reference levels. The RLDs would accompany other instruments in the CEFR "toolkit" (see Figure 1), including, for example, a manual for aligning examinations to the framework.
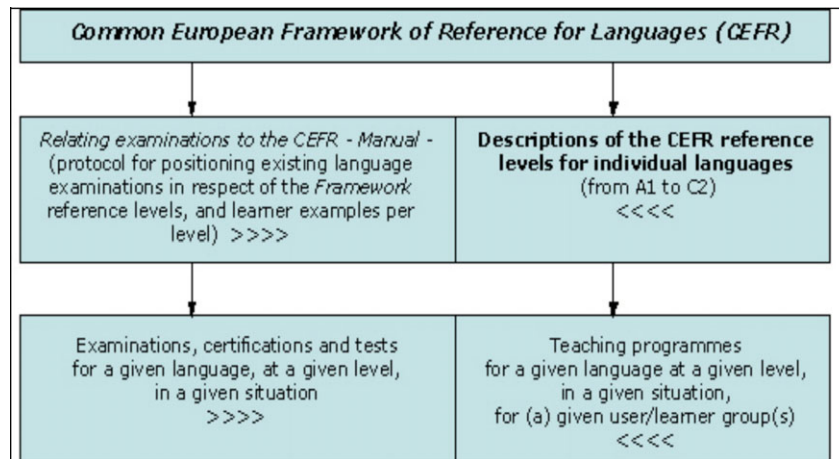
Figure 1   Diagram showing the relationship between Council of Europe reference instruments for language teaching and learning (the so called CEFR "toolkit"). Taken from Council of Europe website.

To assist national teams in developing RLDs for their own national or regional languages, the Language Policy Division developed a *Guide for the production of RLDs*. A seminar was held in Strasbourg in December 2005 for teams reporting on RLD projects that were already being developed or were being planned. The RLDs for German, *Profile deutsch A1, A2, B1, B2*, had already been published as a result of an international collaborative project, and were presented in detail at this seminar. In total, projects representing seventeen languages were presented, including English Profile for establishing a set of RLDs for English proposed by Nick Saville.

English Profile had first been discussed as a potential RLD project for English in June 2005, when representatives from six organisations working in related fields met in Cambridge. Participating were University of Cambridge ESOL Examinations; Cambridge University Press; University of Cambridge Research Centre for English and Applied Linguistics (RCEAL) with the University's Computer Laboratory; the British Council; English UK, and the University of Bedfordshire Centre for Research in English Language Learning and Assessment (CRELLA). An initial three-year English Profile project plan funded by the participating organisations was agreed.

The approach adopted for the *Profile deutsch* project was influential at the start, not least in the choice of the name *English Profile*. Already at this early stage it was intended that the English RLDs would include innovative features and emphasise empirical research approaches rooted

in data. From the outset, the diversity of the English Profile core participant group was held to be one of its main strengths. At early meetings, working groups were set up to focus on key research strands such as corpus linguistics, language pedagogy, curriculum and assessment. These elements have remained part of English Profile as demonstrated below.

In the period that followed the 2005 Strasbourg meeting, English Profile was formally established with a project coordinator in Cambridge. The British Ambassador to the Council of Europe was asked to propose the project as an officially recognised RLD project for the English language. Since then regular progress reports have been submitted to the Language Policy Division through John Trim, who has acted as an observer on the division's behalf, as well as playing a key advisory role in the development of English Profile.

3. Key features of English Profile

At the outset, the objectives of English Profile (as stated, for example, by Saville, IATEFL, 2006) were as follows:

> To provide descriptions of English covering all six levels of the CEFR in one coherent approach through a programme of interdisciplinary research that would be informed by theories from psycholinguistics and second language acquisition and would be based on empirical evidence especially from learner corpora. The outcomes would include specifications of learning objectives by level (A1 to C2) and provide the basis of detailed diagnostics by level – grammatical, lexical and functional exponents (later to be known as criterial features). Dissemination would be through the use of interactive (web-based) technology which would make the descriptions and data covering the six CEFR levels publicly available and provide additional support for users through illustrative examples of learner performance by level.

These English Profile objectives still stand. They still characterise the focus of the programme; the nature and theoretical heritage of the research; the intended constructs and characteristic of the outcomes, and the modes of their delivery and follow up.

The English Profile descriptions of English language reference levels will be described in terms of the features of English which learners attempt to learn as they progress from beginner level (A1) to expert user level of the language (C2). The units of currency for English Profile descriptions of the language levels will be grammatical, lexical and functional exponents derived empirically as criterial for the levels concerned, that is serving as a basis for distinguishing one proficiency level from another. The research carried out by English Profile partners will indeed test hypotheses derived with reference to appropriate linguistic theory and applications. The data for English Profile research will, eventually and in the main, be authentic corpora of the written and spoken performance of non-native users of English at each of the A1 to C2 reference levels. English Profile information sharing is indeed interactive and web-based among programme members, both founding members and more recent, and to the general public as well.

It is likely that English Profile will also prove a useful source of learning objectives for English language curriculum and syllabus development. The English Profile RLDs will focus on the features of English that are present in the language that learners themselves are able to produce (in their speech and writing) at all six reference levels. The Profile should also be of

value in developing learning-oriented assessments and in helping teachers identify the specific strengths and weaknesses of their learners at different stages of their learning programmes.

This article will focus on all these elements of English Profile, as well as key English Profile constructs, data sources, modes of analysis and the dissemination of ideas and findings.

The English Profile is clearly very ambitious. Given the data-based research rigour demanded, its aims cannot be achieved within a short time or without an effective project-based structure. In order to provide this, English Profile goals were subdivided into objectives which could be achieved within a three-year timeframe, known as the English Profile project phase (2005 to 2008). At the end of this period progress was reviewed and it was agreed to extend the collaboration into a rolling programme of research from 2008–9 onwards (see summary in Appendix 1 below).

A starting point for the project team was to review the 'general principles and approach' adopted by the CEFR, especially its six reference levels and its underlying socio-cognitive model of language use covering the five language skills (reading, writing, listening, spoken interaction and spoken production) and the four principal domains of language use (personal, public, occupational and educational). It was also agreed, in the first instance, to focus (though by no means exclusively) on the most underspecified levels of the CEFR, namely the C-levels (C1 and C2) and the A1 level (which the T-series did not cover).

4. The socio-cognitive approach

The socio-cognitive approach adopted by the authors of the CEFR is succinctly summarised in Chapter 2 of the 2001 volume (see above):

> Language use, embracing language learning comprises the actions performed by persons who as individuals and as social agents develop a range of *competences*, both *general* and in particular *communicative language competences*. They draw on the competences at their disposal in various contexts under various *conditions* and *constraints* to engage in *language activities* involving *language processes* to produce and/or receive texts in relation to *themes* in specific *domains*, activating those *strategies* which seem most appropriate for carrying out the *tasks* to be accomplished. The monitoring of these actions by the participants leads to the reinforcement or modification of their competences.

> (Cambridge University Press, 2001: 9, emphasis in original)

This is also referred to as "an action-oriented approach" in so far as learners are viewed as "social agents" who operate in contexts in which "communicative acts" are embedded. Jones (2009: 2–8) represents this as a "triangle" of variables, as shown in Figure 2.

The socio-cognitive model which this "triangle" represents has also become increasingly important as a way of addressing construct validity in language assessment (see, for example, Weir, 2005). The model has been adopted by Cambridge ESOL as the basis for the validation of Cambridge ESOL examinations (Shaw & Weir, 2007; Khalifa & Weir, 2009) which have provided significant corpus data being used in the English Profile research programme (see below).

The focus on *context* (both language *use* and language *learning* contexts) became a major concern early on in the English Profile Programme and it was determined that the English
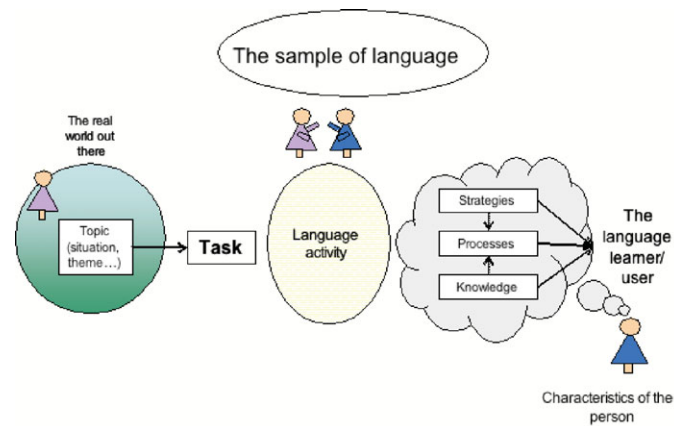
Figure 2 Triangle of variables: learner, task, performance (sample of language).

Profile research teams would ensure the collection of data to reflect how English is learned and used around the world (see *Setting up of the English Profile Network* below).

5. The Cambridge Learner Corpus and the use of learner data

In order to empirically investigate learner language at each of the CEFR levels, large samples of writing and speech are clearly needed. A long-term aim of English Profile is therefore to build extensive and representative language corpora, useful for both the investigation of English Profile hypotheses about language levels and for ideas on their learning/teaching and assessment. Since at the outset of the Programme such corpora were not yet available to the researchers, it was decided that in the first phase of English Profile existing data, in particular the Cambridge Learner Corpus (CLC), would be used.

The CLC, collected since the early 1990s in a joint project between Cambridge University Press and Cambridge ESOL, consists of written English produced by candidates who took Cambridge ESOL examinations at one of the CEFR levels – A2 to C2 (see CLC website). By December 2009 the CLC consisted of 37 million words of learner written English, of which more than 19 million words are error-coded (www.cup.es/elt/corpus/learner_corpus.htm). The data in the CLC include extended responses to tasks which require the candidates to produce written text as opposed to one word or short answers (as in cloze-type tasks). Along with the candidates' written responses, the corpus also contains meta-data about candidate gender, age, first language, reason for taking the exam. Also available is evidence of candidates' proficiency in English: overall grade and marks on the other exam components (typically reading, listening and speaking). Item-level information from these components can be accessed, although this is not stored in the CLC database and so is not easily retrievable. The tasks to which the candidates have responded may be accessed, both as images and as a searchable sub-corpus.

At present, English Profile researchers have mainly been using data from a subset of the CLC, amounting to some 26 million words, more than half of which are *error-coded*. There
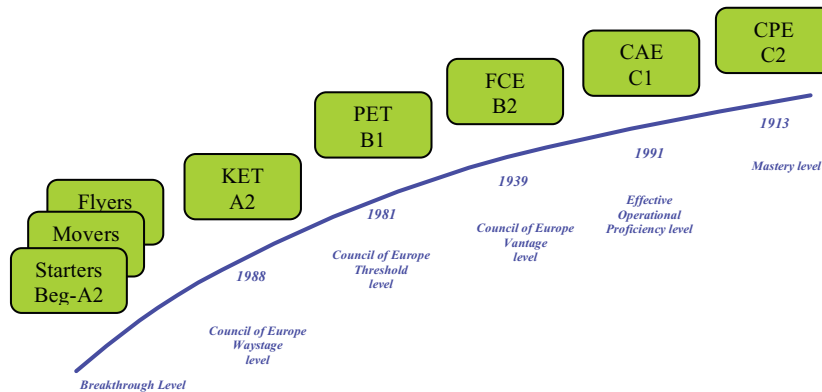
Figure 3 Alignment of Cambridge ESOL exams with CEFR levels.

Table 1   *CLC error codes, explanations and exemplifications.*

Error Code	Explanation	Exemplification
RN	Replace noun	Have a good travel (journey)
RV	Replace verb	I existed last weekend in London (spent)
MD	Missing determiner	I spoke to President (the)
AGN	Noun agreement error	One of my friend (friends)
AGV	Verb agreement error	The three birds is singing (are)

are two main types of examination in this corpus subset. The Main Suite is a set of general purpose exams consisting of the Certificate of Proficiency in English (CPE), Certificate in Advanced English (CAE), First Certificate in English (FCE), Preliminary English Test (PET), and Key English Test (KET). The examinations were developed alongside the CEFR and have converged with its reference levels (see Jones, 2000, 2001, 2002; Taylor & Jones, 2006 for discussions of the alignment of exams with the CEFR levels). The alignment is shown in Figure 3.

The Cambridge ESOL purpose-specific Business English Certificates (a professional suite that focuses on business English) consist of BEC Higher (C1), BEC Vantage (B2), and BEC Preliminary (B1).

As part of English Profile, the *coding* and *parsing* of elements in the CLC have extended the kinds of analysis which can be carried out and allow the research teams to investigate a much wider range of English language features. The CLC's system consists of over 70 two-part error codes, specifying the type of error and the part of speech it applies to (see Nicholls, 2003 for a description of the error-coding system). Table 1 provides examples.

Relevant to English Profile's need for empirical evidence of key features of English language use at the CEFR levels, these error codes, together with candidate and exam meta-data, already enable the calculation of frequency statistics for each exam level, language group, age and so on.

To enable searches beyond the individual word level and over a wide range of lexical and grammatical features, a *computational* strand of research was introduced into the English Profile Programme at an early stage. The CLC was tagged for parts-of-speech and parsed using the Robust Accurate Statistical Parser (RASP) by researchers at English Profile partner University of Cambridge Computer Laboratory under the supervision of Professor Ted Briscoe (Briscoe, Carrol & Watson, 2006). RASP is an automatic parsing system incorporating both grammatical information and statistical patterns, and its operation is summarised by Hawkins and Buttery (2009b: 6). More details and examples on the annotation of the CLC with part-of-speech tags, word lemmas, grammatical relations and complexity metrics are to be found in Hawkins & Buttery, 2009b; pages 15–20, and elsewhere in this journal.

It is clear that, in its research to specify key grammatical, lexical, functional and other English language learning features for each CEFR level, English Profile is likely to raise a number of relevant second language learning (SLA) issues. Previous research in SLA has been constrained by the lack of consistent assessment-based classification of learners according to their second language proficiency, through language databases large enough to permit correlations across diverse lexical and grammatical features; justify conclusions about the interrelation of developing SLA features; or to allow the study of L1 transfer effects across all major language families, as represented in the CLC data from learners from 130 first languages. As indicated below and elsewhere in this journal, the cross-linguistic differences at each CEFR level are now a major focus in English Profile research.

The emerging performance patterns per CEFR level are potentially informative for our understanding of SLA, for example the order of acquisition of linguistic features and the interaction of factors such as frequency, complexity and L1 transfer. Related questions which English Profile researchers are beginning to address include:

- How do the different kinds of criterial features (lexical semantic, morpho-syntactic, syntactic, discourse, notional, functional, etc.) interrelate?
- To what extent does the criteriality of features vary depending on the L1 of the learner?
- Which criterial features can be used as diagnostics at the individual learner level?
- What is the effect of task type on learner production and criterial features? (Parodi, 2008)
- How does the type of context in which spatial verbs occur help explain the spatial information findings (Hendriks, 2008)?

and in particular

- Which linguistic features realise which language functions across the CEFR levels?

Answering these questions is fundamental for SLA knowledge. English Profile research into these questions also tends to help knit together the different strands of the English Profile research programme.

6. Building new corpora and extending the skills

The immediate future of English Profile (2010) will involve extending the current analyses to broader samples from the CLC and collecting other kinds of written data from learners

of English worldwide. A major data collection exercise undertaken from 2008 onwards is already bringing into the Programme new corpora of learner written English at various levels.

A significant challenge being addressed is how to include *spoken language* in the analysis of learner English language at each CEFR level (McCarthy & Saville, 2009). Such a profile will complement the current linguistic profile being investigated for writing. Professor Michael McCarthy refers elsewhere in this journal to his 160,000 words of ESOL oral exam data and to the Cambridge, Limerick and Shannon (CLAS) Corpus (see also below) of 120 hours of recordings made in a variety of pedagogic contexts, which, when transcribed, could approach the target of a million words of learner oral English. Further spoken English corpora from other research projects continue to be forthcoming for possible English Profile use, using developing transcription conventions and search methods.

In terms of the socio-cognitive model (above) and the importance it attaches to context, the building of new corpora explicitly targets learners with specific characteristics (age, L1, etc.) who are learning English in different pedagogic environments around the world (in school classrooms, at universities, through self study programmes, online, etc.) and who are required to produce specific language samples based on a wide range of different prompt material (tasks).

7. Setting up the English Profile network

One of the key qualities which had recommended English Profile to the Council of Europe at the outset had been the intention to ensure that the project was an international one, avoiding a purely Anglo-centric approach. This was clearly important to reflect the unique position of English as a global language. It was a stated aim of English Profile from the start to involve English language learning, teaching and assessment stakeholders around the world. Thus, the English Profile February 2009 seminar was attended, in addition to participants from the main research partners (namely Cambridge University Research Centre for English and Applied Linguistics (RCEAL), and Computer Laboratory, Cambridge University Press, Cambridge ESOL Exams, the Centre for Research in English Language Learning and Assessment (CRELLA), University of Bedfordshire) by representative from MATSDA UK, Computense University of Madrid, the University of Applied Sciences and English Expertise in Cologne, and the University of Vienna. More recent non-UK involvements in English Profile have included the establishment in October 2009 of the Cambridge, Limerick and Shannon (CLAS) Corpus (see above), the result of a year-long collaboration with the University of Limerick's Mary Immaculate College (MIC), and the Shannon College of Hotel Management in Ireland.

Because of the international framing of English Profile as a programme, and also the diversity of the approaches within it, members have been able to take part in the network according to their interests and abilities. Some are now forming regional networks to focus on matters relating to English language learners in a specific region or with a specific linguistic background (e.g. in Japan). Such regional English Profile Networks, established in a number of countries, should provide a focus for discussion of key issues relating to the research questions which English Profile is investigating and eventually become model platforms for curriculum innovation and reform. While some choose to become part of the network based

on geographic considerations, others take a more thematic interest and contribute to the research strands according to their own interests and resources.

As the projects have developed, it has become increasingly important for more data to be collected and developed into corpora. Network members now play a key role in collecting these data which will be critical to delivering the empirically measurable results and conclusions which are envisaged.

From its outset, English Profile was intended to be a partnership based on collaboration and with an inclusive agenda. As the project has moved from scoping and planning its work into the research projects that will lead to publishable results, the importance of bringing in further partners has grown. The English Profile Network has now become an extended base of academics, government advisors and educationalists with an interest in the English language.

8. Three key English Profile research projects

Given the objectives of English Profile, to develop a detailed profile of the English language of learners from a range of L1 backgrounds at the reference levels specified in the CEFR, it is logical that three of the main research teams working to this end should focus on

1. The grammatical features
2. The words and their meanings
3. The communicative language functions

the knowledge and use of which are evidenced as distinguishing the CEFR levels.

The contributions in this journal by Filipovic & Hawkins; Capel, and Green (see below) illustrate the constructs, research data and analyses, and some of the early findings in these areas, towards English Profile. As suggested above, the outcomes from these three English Profile projects will include, or signpost, specifications of characteristic English language learning abilities by CEFR level in terms of detailed grammatical, lexical and functional exponents which are found to be *criterial* to the level concerned. These specifications, again as suggested in other articles in this journal, will also inform and enhance a key future English Profile aim, probably including the data and findings from existing criterial feature research, that will make it possible to foster a closer relationship between English Profile outcomes and teachers and learners of English in their different contexts worldwide (Alexopoulou, 2008).

9. English Profile publications and dissemination plan 2010 onwards

It is appropriate that the English Profile team returned to Harrogate for IATEFL 2010, exactly four years after the official UK launch event. At the conference, John Hawkins discussed the outcomes of the corpus-based strand, exemplifying this with illustrations from the Hawkins and Filipovic English Profile volume (see below). He also explained for a teaching-focused audience how the criterial features which are emerging from English Profile research may be helpful in a number of practical ways with benefits for them.

At IATEFL 2010, in addition to stressing the open and participatory nature of the English Profile network, the emphasis was on the achievements and outcomes to date, including the publishing agenda and precise information about the first two volumes in English Profile studies series, to be published by Cambridge University Press in 2011, on English Profile grammatical criterial features (see above) and the theory and practice of language construct definition across the ability range. The English Profile Wordlists (see Capel in this journal) are also being published online by Cambridge University Press in 2011.

A key objective in 2010 will be the development of a longer-term publishing plan to incorporate both the web-publishing initiatives (English Profile Journal, English Profile Wordlists etc.) and the book series. Since January 2010, an editorial group has met on a regular basis to ensure that the plan for the next five years is fleshed out. Once in place this can also be used to inform research priorities. The plan will include detailed proposals for the subsequent editions of currently planned volumes as well as for the publication of new work on other strands (pedagogy, assessment, the role of the teacher, materials development etc.).

It is also planned to set up a formal English Profile association to strengthen the English Profile network and to formalise the ways in which network members can take a fuller part in English Profile activities, including contributions to the funding and management of the programme. The new association would enable other commercially-oriented organisations to play an active role in future developments.

Given the nature of English Profile as a network with participants from different disciplines, it is likely that some key terms and concepts will be interpreted differently by different researchers. In order to facilitate communication and to enhance coherence across the publications under preparation, an English Profile glossary is to be developed, online and open-ended to maximise participation of members across English Profile research fields.

Conclusion

In this article we have attempted to summarise key features of English Profile Programme, a major inter-disciplinary, research-based collaborative project to develop detailed reference level descriptions (RLDs) of the English of L2 language learners linked to the general principles and approaches of the Council of Europe. The further articles in this journal describe in detail English Profile research in the key areas selected to fulfil this ambitious aim.

Appendix 1 – Summary of English Profile activities: 2005–9

2005–6 Setting up the project

- Appointed Cambridge ESOL-based Project Coordinator
- Upgraded the accessibility of the Cambridge Learner Corpus (CLC)
- Hypothesised criterial features defining each proficiency level
- Prepared a draft research agenda identifying relevant research areas
- Organized a seminar for all the partners and collaborators involved

- Launched the project at IATEFL Conference in Harrogate (April 2006)
- Registered with the Council of Europe as an RLD project for English

2006–7 Extending the research agenda

- Held a number of seminars, presentations and other events
- Started setting up a network of external collaborators around the world
- Continued work on research projects
- Began research in new areas
- Developed the website further, adding password protected areas

2007–8 Building a network

- Held seminars and events to extend the network of collaborators around the world
- Continued work on research projects
- Published the special *English Profile* edition of Research Notes
- Started work on *English Profile Wordlists* project

2008–9 Consolidatation – from *project* to *programme*

- Launched new website with news items, log-in, corpus information
- Launched preview version of *English Profile Wordlists* on English Profile website
- Conducted pilot studies for corpus data collection in Italy, Japan and Spain
- Set up web interface for collection of written corpus data (funded by ESOL)
- Started collection of spoken corpus data in Shannon, Ireland
- Widened network of researchers
- Funded two Cambridge Assessment research scholarships (ESOL)
- Commissioned two English Profile Studies volumes: Criterial Features and Functional Progression
- Gained approval from Cambridge University Press syndics for publication of *English Profile Journal*, an open-access online journal
- Gained approval from Council of Europe to convert T-series to ebooks
- Published Trim *Breakthrough* manuscript online
- Won EU Lifelong Learning grant to support Network activities

References and further reading

Alexopoulou, T. (2008). Building new corpora for English Profile. *Research Notes*, *33*, 15–19.

Briscoe, E., Carroll, R. & Watson, R. (2006). The second release of the RASP system. In *Proceedings of the COLING/ACL 2006 Interactive Presentation Sessions*, Sydney, Australia. Available online at http://acl.ldc.upenn.edu/P/P06/P06-4020.pdf.

Capel, A. (2009). A1–B2 vocabulary: Insights and issues arising from the *English Profile Wordlists* project. Paper presented at the English Profile Seminar, Cambridge, 5–6 February 2009.

Coste, D. (2007). Contextualising the use of the CEFR for Languages. Paper presented at the Council of Europe Policy Forum, Strasbourg 2007. Available online at http://www.coe.int/

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

Doughty, C. J. & Long, M. H. (eds.) (2005). *The Handbook of Second Language Acquisition*. Oxford University Press.

Green, T. (2008). English Profile: Functional progression in materials for ELT. *Research Notes*, *33*, 19–25.

Filipović, L. (2009). English Profile report. Internal Cambridge ESOL report, July 2009.

Filipović, L. & Hawkins, J. (2009). Emerging patterns and principles in the Cambridge Learner Corpus of English. Paper presented at the ELLSIIR conference, Belrgade, Serbia, 4–6 December 2009.

Hawkins, J. A. & Filipović, L. (in preparation). Criterial features of English across the reference levels of the CEFR, unpublished manuscript.

Hawkins, J. A. (2009). An efficiency theory of complexity and related phenomena. In D. Gill, G. Sampson & P. Trudgill (eds.) *Complexity as an evolving variable*. Oxford University Press.

Hawkins, J. A. (2004). *Efficiency and complexity in grammars*. Oxford University Press.

Hawkins, J. A. (1994). *A performance theory of order and constituency*. Cambridge University Press.

Hawkins, J. A. & Buttery, P. (in preparation). Using learner language from corpora to profile levels of proficiency: Insights from the English Profile Programme. *English Profile Journal*.

Hawkins, J. A. & Buttery, P. (2009). Criterial features in learner corpora: Theory and illustrations. Paper presented at the English Profile Seminar, Cambridge, 5–6 February 2009.

Hendriks, H. (2008). Presenting the English Profile Programme: In search of criterial features. *Research Notes*, *33*, 7–10.

Hulstijn, J. H. & Schoonen, R. (2010). Developmental stages in second-language acquisition and levels of second-language proficiency: Are there links between them? *English Profile Journal*. http://journals.cambridge.org/action/displayJournal?jid=EPJ

Jones, N. (2000). Background to the validation of the ALTE Can Do Project and the revised Common European Framework. *Research Notes*, *2*, 11–13.

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

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. 167–83). Strasbourg: Council of Europe. Available online at www.coe.int/T/DG4/Portfolio/documents/case_studies_CEF.doc.

Jones, N. (2009). The classroom and the Common European Framework: Towards a model for formative assessment, *Research Notes*, *36*, 2–8.

Khalifa, H. & Weir, C. (2009). *Examining Reading: Research and practice in assessing second language reading*. University of Cambridge ESPOL Examinations and Cambridge University Press.

Kurteš, S. & Saville, N. (2008). The English Profile Programme: An overview. *Research Notes*, *33*, 2–4.

Little, D. (2006). The Common European Framework of Reference for Languages: Contents, purpose, origin, reception and impact. *Language Teaching*, *39*.3, 167–90.

Little, D. (2007). The Common European Framework of Reference for Languages: Perspectives on the making of supranational language education policy. *Modern Language Journal*, *91*, 645–55.

McCarthy, M. & Saville, N. (2009). Profiling English in the real world: What learners and teachers can tell us about what they know. Paper presented at American Association for Applied Linguistics Conference, Denver, Colorado, 21–24 March 2009.

Milanovic, M. (2009). Cambridge ESOL and the CEFR. *Research Notes*, *37*, 2–5.

Nicholls, D. (2003). The Cambridge Learner Corpus: Error coding and analysis for lexicography and ELT. Available online at http://ucrel.lancs.ac.uk/publications/CL2003/papers/nicholls.pdf.

Odlin, T. (2005). Cross-linguistic influence. In C. J. Doughty & M. H. Long (eds.), pp. 436–89.

Parodi, T. (2008). L2 morpho-syntax and learner strategies. Paper presented at the Cambridge Institute for Language Research Seminar, Cambridge, 8 December 2008.

Salamoura, A. (2008). Aligning English Profile research data to the CEFR. *Research Notes*, *33*, 5–7.

Salamoura, A. & Saville, N. (2009). Criterial features across the CEFR levels: Evidence from the English Profile Programme. *Research Notes*, *37*, 34–40.

Schachter, J. (1974). An error in error analysis. *Language Learning*, *24*, 205–14.

Shaw, S. & Weir, C. (2007). *Examining Writing: Research and practice in assessing second language writing*. University of Cambridge ESOL Examinations and Cambridge University Press.

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

Taylor, L. & Jones, N. (2006). Cambridge ESOL exams and the Common European Framework of Reference (CEFR). *Research Notes*, *24*, 2–5.

Trim, J. L. M. (2009). *Breakthrough*. Available online at www.englishprofile.org.

van Ek, J. & Trim, J. L. M. (1990a/1998a). *Threshold 1990*. Cambridge University Press.

van Ek, J. & Trim, J. L. M. (1990b/1998b). *Waystage 1990*. Cambridge University Press.

van Ek, J. & Trim, J. L. M. (2001). *Vantage*. Cambridge University Press.

Weir, C. (2005). Language testing and validation: An evidence-based approach. Basingstoke: Palgrave Macmillan.

Williams, C. (2007). A preliminary study into the verbal subcategorisation frame: Usage in the CLC. RCEAL, Cambridge University, unpublished manuscript.