EnglishProfile

Introducing the CEFR for English

Version 1.1

August 2011

English Profile
Introducing the CEFR for English

This booklet is aimed at ELT professionals, including curriculum planners, materials writers and teachers. It will help you make decisions about which English language points are suitable for learning, teaching or assessing at each level of the Common European Framework of Reference (CEFR; Council of Europe 2001).

Section 1 explains what English Profile is whilst sections 2, 3 and 4 describe how Grammar, Functions and Vocabulary are being researched in English Profile, together with a summary of the latest findings in all three research strands. Section 5 suggests how you can use these resources and section 6 describes where the information in English Profile comes from and how reliable it is. Section 7 explains where you can access more information and resources and how you can get involved with the ongoing development of English Profile, followed by the key references for English Profile research to date and a list of current English Profile Network members.

Contents

1 What is English Profile?

The aims of English Profile are:

- To set up and manage a collaborative programme of interdisciplinary research to produce *Reference Level Descriptions for English* linked to the general principles and approaches of CEFR.
- To provide a core set of reference tools for practitioners working in English language education.

English Profile is a long-term, collaborative programme of interdisciplinary research whose goal is to transpose the Common European Framework of Reference for the English language (CEFR; Council of Europe 2001) and for teaching and assessment purposes where English is the language being learned. The intended output is a 'profile' of English language learners in terms of the six proficiency bands of the CEFR- A1 to C2 (see Saville & Hawkey 2010). The *English Profile Programme* will do this by providing Reference Level Descriptions for English to accompany the CEFR. These descriptions cover what learners know and can do in English at each of the six CEFR levels. English Profile is registered with the Council of Europe and is managed by a core group of collaborators at the University of Cambridge.

The research being carried out at the heart of the English Profile Programme is innovative, providing measurable, evidence-based answers to important questions about how people acquire English and how they can improve their skills. As well as adding to our understanding of language learning, the English Profile Programme is producing practical outcomes that can be used in the development of curricula, course materials, teaching guides and assessment systems. This publication traces progress and outcomes in three main current areas of research for English Profile: the grammatical, functional and lexical features of learner English. But English Profile will also describe learner English at each CEFR level in other linguistic areas, including aspects of spoken language such as pronunciation.

An innovative feature of English Profile, distinguishing it from previous work in this field, is that research is based on electronic corpora of learner data, including the largest annotated corpus of English language learner test output in the world: the Cambridge Learner Corpus. This approach is producing results which can be empirically measured and which are not predictable from current language learning theories alone. Researchers are also starting to focus on the impact of different first languages, learning contexts and the effects of language transfer on learning at the different CEFR levels (A1 to C2).

A steadily growing number of academics, government advisors and educationalists make up the *English Profile Network*. Network Partners contribute directly to the development of English Profile by providing access to data or contributing to work in progress through participation in workshops and seminars.

In summary, English Profile provides essential information for curriculum planners, teachers, materials writers, test developers and researchers. The English Profile Programme aims to provide these ELT professionals with resources, information and events, including:

- English Vocabulary Profile (EVP) – a rich online vocabulary database by CEFR level
- English Grammar Profile – a database of grammatical structures by CEFR level (under development)
- English Functions Profile – a database of real English examples for various functions in different contexts by CEFR level (under development)
- English Profile Glossary – an online glossary including concise definitions of key EP terminology
- English Profile Journal – on online peer-reviewed journal for EP-related research
- The English Profile Studies series – launched in 2011 this series is dedicated to reporting different aspects of research and development related to the EP Programme
- Word of the Week email updates based on the EVP
- English Profile Network community website, **www.englishprofile.org**

- Presentations at international education, applied linguistics and language testing conferences, e.g. IATEFL, AILA, LTRC
- Regular EP Research Seminars (annually in Cambridge), EP Network Seminars (twice a year outside the UK) and other workshops.

For the latest information about English Profile and news of future events, workshops and publications, see **www.englishprofile.org**

Who has developed English Profile?

The founding partners are:
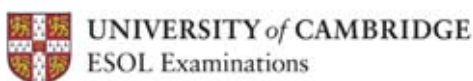
Research is led by Cambridge ESOL and Cambridge University Press, with contributions from the Research Centre for English and Applied Linguistics and the Cambridge Computer Laboratory.[1]

Cambridge University Press and Cambridge ESOL are the main funding partners in English Profile. In addition, English Profile has a growing number of Network Partners, including universities, schools, language training centres and government departments, as well as individual researchers and specialists (see Section 9).

The development of the CEFR

The Common European Framework of Reference for Languages is a common framework of language ability divided into six main levels ranging from beginner to advanced. It appeared in its published form in 2001, ten years after the Rüschlikon Conference of 1991 which concluded that a "common framework of reference" of this kind would be useful as a planning tool to promote "transparency and coherence" in language education. In the decade since its publication this ambition has been achieved to a large extent and the document itself has been translated into 37 languages, widely disseminated in Europe and in parts of Asia and Latin America (see Little 2007 for an overview). It is important to remember, however, that the CEFR in that format was intended to be "a work in progress" rather than the finished article.

The CEFR was therefore envisaged as a planning tool which could provide a "common language" for describing objectives, methods and assessment in language teaching, as put into practice in diverse contexts for many different languages. It was to facilitate the development of syllabuses, examinations, textbooks and teacher training programmes, and in particular, to stimulate reflection and discussion. As the CEFR authors themselves emphasise in their *Notes for the user*:

> We have NOT set out to tell practitioners what to do or how to do it. We are raising questions not answering them. It is not the function of the CEF to lay down the objectives that users should pursue or the methods they should employ. (Council of Europe 2001: xi)

---

[1] Part of the Department of Theoretical and Applied Linguistics from August 2011.

The six reference levels have been particularly influential and have generated a great deal of discussion in the fields of curriculum development, language teaching, and especially in assessment (see Coste 2007). The levels are described through the six-level *Global Scale* (A1 to C2) and the *Illustrative Descriptors* that can be applied to the learning and teaching of *any* language. Table 1 presents the Global Scale descriptors for the six main CEFR levels, showing how the CEFR is a general document that needs to be further specified and contextualised for each area of use.

Table 1: Global Scale descriptors for CEFR levels (Council of Europe 2001: 24)

Proficient User	C2	Can understand with ease virtually everything heard or read. Can summarise information from different spoken and written sources, reconstructing arguments and accounts in a coherent presentation. Can express him/herself spontaneously, very fluently and precisely, differentiating finer shades of meaning even in more complex situations.
	C1	Can understand a wide range of demanding, longer texts, and recognise implicit meaning. Can express him/herself fluently and spontaneously without much obvious searching for expressions. Can use language flexibly and effectively for social, academic and professional purposes. Can produce clear, well-structured, detailed text on complex subjects, showing controlled use of organisational patterns, connectors and cohesive devices.
Independent User	B2	Can understand the main ideas of complex text on both concrete and abstract topics, including technical discussions in his/her field of specialisation. Can interact with a degree of fluency and spontaneity that makes regular interaction with native speakers quite possible without strain for either party. Can produce clear, detailed text on a wide range of subjects and explain a viewpoint on a topical issue giving the advantages and disadvantages of various options.
	B1	Can understand the main points of clear standard input on familiar matters regularly encountered in work, school, leisure, etc. Can deal with most situations likely to arise whilst travelling in an area where the language is spoken. Can produce simple connected text on topics, which are familiar, or of personal interest. Can describe experiences and events, dreams, hopes & ambitions and briefly give reasons and explanations for opinions and plans.
Basic User	A2	Can understand sentences and frequently used expressions related to areas of most immediate relevance (e.g. very basic personal and family information, shopping, local geography, employment). Can communicate in simple and routine tasks requiring a simple and direct exchange of information on familiar and routine matters. Can describe in simple terms aspects of his/her background, immediate environment and matters in areas of immediate need.
	A1	Can understand and use familiar everyday expressions and very basic phrases aimed at the satisfaction of needs of a concrete type. Can introduce him/herself and others and can ask and answer questions about personal details such as where he/she lives, people he/she knows and things he/she has. Can interact in a simple way provided the other person talks slowly and clearly and is prepared to help.

The CEFR reference levels and illustrative descriptors (exemplified in Table 1) were intended to be used for the organisation of learning and teaching within educational systems. These levels and descriptors were to provide a communication tool to assist ELT practitioners in practical ways, having been selected and synthesised from existing scales which had been developed and operationalised in many diverse contexts. The CEFR itself, representing forty or more years of work by the Council of Europe Modern Languages Division, builds on earlier studies of levels of language competences such as Threshold Level (van Ek 1975; van Ek & Trim 1998b), Waystage and Vantage Levels (van Ek & Trim 1998a; 2001).

However, although the CEFR is an intuitively helpful descriptive scheme for researchers, curriculum designers, teachers, teacher trainers, and language testers, in many cases the existing scales and related descriptors have not proved to be operationally adequate as they stand. The details are not specific enough to help these professionals make decisions about language teaching and testing. So, while the CEFR can act as a focal point for reference purposes, it must remain open to further development. In other words, the CEFR is not the finished article but needs to be adapted or developed further for each specific context in which it is to be used. John Trim, one of the "fathers" of the CEFR, and now Council of Europe observer to the English Profile Programme and an active member of the growing EP team, summarises this situation succinctly:

> Overall, the apparatus of level description in the CEFR is rich and well differentiated for different purposes and users. Even so, experience over the past decade has shown that for high stakes purposes, particularly the valid and reliable calibration of qualifications and the tests and examinations leading to their award, the CEFR cannot be used as a 'stand alone' document. Indeed, it is probably impossible for any such document to be so used. (preface to Green 2011: xi)

Importantly the CEFR is neutral with respect to the language being learned. This means that the users have to decide what actually gets taught or assessed in terms of the linguistic features of a specific language at each of the common reference levels. To ensure that the framework is used appropriately and can be adapted to local contexts and purposes, the Council of Europe has encouraged the production of instruments and support materials to complement the CEFR. These instruments (sometimes known as the CEFR toolkit) include *Reference Level Descriptions (RLDs) for national and regional languages*. RLDs seek to provide language-specific guidance for users of the Framework; the aim is to "transpose" the Framework descriptors that characterise the competences of users or learners at a given level into the linguistic material which is specific to a given language (i.e. grammar, lexical items etc.) and considered necessary for the implementation of those competences. In providing a description of the language across all six levels, the grammatical and lexical progression which is central to the learning of that language can be addressed more precisely within the Framework concept. The RLDs represent a new generation of descriptions which identify the specific forms of any given language (words, grammar, etc.) at each of the six reference levels which can be set as objectives for learning or used to establish whether a user has attained the level of proficiency in question.

To assist the teams in developing RLDs for their own languages, the Language Policy Division of the Council of Europe produced a general *Guide for the production of RLDs* which was discussed at a seminar held in Strasbourg in December 2005 (see Council of Europe website for details). Projects representing seventeen languages were presented, including a proposal for English put forward by the University of Cambridge (represented at the meeting by Cambridge ESOL). This proposal subsequently became known as the English Profile Project, which, in turn, became the English Profile Programme in 2008.

The English Profile Programme – the CEFR for English

A major objective of the EP Programme is to analyse language produced by learners of English in order to throw light on what they *can* and *can't* do with the language at each of the Common European Framework of Reference levels, for example, in using the grammar and lexis at their disposal. The founder members of the EPP first met in Cambridge in mid-2005 to discuss the possibility of setting up an RLD project for English. Participating in those discussions were several departments of the University of Cambridge (Cambridge ESOL, Cambridge University Press, the Research Centre for English and Applied Linguistics and the Computer Laboratory), together with representatives from the British Council, English UK, and the University of Bedfordshire (Centre for Research in English Language Learning and Assessment).

As a result of those meetings, the English Profile Project was established by a core group of partner organisations in 2005 to with the aim of making the CEFR more specifically explicit with regards to English language learning, teaching and assessment. The core group was extended to create an English Profile Network from 2008 onwards (see pages 63-64 for a list of EP Network Partners). Coordination of the Programme is based at the University of Cambridge and involves interdisciplinary collaboration between different departments.

From an early stage the English RLDs were intended to be innovative with an emphasis on empirical research rooted in data (such as learner corpora), and the collection of representative samples of learner language which could be used to explore language development across the reference levels. This has involved collaborators from the EP Network in different parts of the world who can supply samples of speaking and writing produced by learners. It is an aspect of the project which has received external funding by the European Commission and is now well underway (see the EU-funded English Profile Network project members on page 64).[2] It has also required technical resources in developing new electronic corpora and analytical techniques so that the samples of learner language can be stored, accessed and analysed effectively.

Given the interdisciplinary nature of the research being undertaken, research teams are engaged in parallel and simultaneous investigations on a set of related research questions, observing them from different angles. This English Profile publication brings together findings from three areas of investigation – grammar, language functions and vocabulary (see Sections 2-4).

"Profile" is a suitable name for the intended outputs of this programme of research. A profile is dictionary-defined as 'a vivid outline of the most outstanding characteristics of the subject' (Webster's). English Profile does indeed seek to specify the "outstanding characteristics" or, in our terms, the *criterial features* validated by empirical research that distinguish each proficiency level of learner English, from A1 to C2. Criterial features are linguistic properties that are distinctive and characteristic of each of the levels. They are not meant to capture all language features that a learner uses at a certain level but to distinguish a level from adjacent higher and lower levels. Criterial features describe changes from one level to the next, and hence are important for both practitioners and theoreticians to know about. This approach is dealt with comprehensively in Hawkins and Filipović (2011) and also in Hawkins and Buttery (2009, 2010) and Salamoura and Saville (2009, 2010).

> **Defining *criteriality:***
>
> "Criteriality" means that the language features concerned serve as a basis for distinguishing one proficiency level from another.
>
> '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'.
> (Saville & Hawkey 2010: 4)
>
> English Profile Glossary

---

[2] The EP Programme is endorsed by the Council of Europe, and has been funded by the European Commission.

The empirical approach makes use of learner data, especially samples of writing and speech produced by learners at different levels of proficiency. The analysis of these data has produced informative results about the language of learners and has added to our understanding of how the grammar and lexis of English is learned by different groups of learners. In addition to a focus on traditional grammatical and lexical features, psycholinguistic factors have also been taken into account, including the effects of language transfer (i.e. the impact of different first languages and learning contexts) on language learning. It is worth noting that some of the results would not always be predicted from language acquisition theories or anticipated by researchers using experience and intuition alone, proving the validity and usefulness of an empirical approach that has real data at its core.

The Cambridge Learner Corpus (CLC) has been central to this work to date. The CLC consists of learners' written English from the Cambridge ESOL examinations covering the ability range from A2 to C2, together with meta-data (including gender, age and first language) and evidence of overall proficiency based on their scores for the other skills components (typically reading, listening and speaking) which currently contains around 45 million words. The table below shows the six CEFR levels A1-C2 and their link to the Cambridge English suite of exams testing General English which have been at the heart of EP research to date.[3]

The demographic information about the learners within the CLC allows researchers to compare different L1 learners with respect to the English that they produce. While lexical analysis has been carried out for many years by researchers in Cambridge ESOL and Cambridge University Press, error coding and parsing of the corpus have extended the kinds of analysis which can be carried out and have allowed the research teams to investigate a wider range of English language features (see Nicholls 2003 on the error-coding system which has informed EP research). A computational strand of research was introduced into the EPP at the outset and the CLC has been tagged and parsed using the Robust Accurate Statistical Parser (RASP) by researchers in the Computer Laboratory under the supervision of Professor Ted Briscoe (Briscoe, Carrol & Watson, 2006). The error coding and the parsing means that sophisticated kinds of grammatical analysis are also possible and are an exciting current focus of research. See the English Grammar Profile section for a fuller description of the CLC.

In addition to describing the "real language" used by learners, English Profile has sought to investigate *the learning dimension* and to connect the empirical work with relevant SLA and linguistic research. In particular EP researchers are interested in "how learners learn English" and how different learning factors interact under various contextual conditions. They have begun to address questions such as:

- How do the different kinds of criterial features (lexical semantic, morpho-syntactic, syntactic, discourse, notional, functional, etc.) interrelate and cluster together to define learner profiles in English? Which linguistic features realise which language functions across the CEFR levels?

---

[3] See **www.CambridgeESOL.org/exams** for a range of exams to meet different needs, including professional and academic purposes.

- How does the profile of the learner vary depending on their L1? What are the pedagogical implications of such L1 effects for the learning, teaching and assessment of English?
- Which criterial features can be used as diagnostics of proficiency at the individual learner level?
- What are the similarities and differences between adult and young learners of English developmentally and at each stage of learning?
- How does learning to speak differ from learning to write/type? What determines communicative success and comprehensibility in these two language modes?
- What is the role of learner and learning strategies?
- How do all the previous factors interact during language learning? How do they predict likely versus less likely patterns of learner output? What type of learning model can accommodate these multi-factor interactions that underpin language learning?

The emerging performance patterns are informative for our understanding of second language acquisition (SLA), e.g. the order of acquisition of linguistic features and the interaction of factors such as frequency and transfer from the first language. Just as EP findings can contribute to new aspects of theory and provide useful insights for developing a model of L2 acquisition, SLA and other linguistic theories inform EP research in this and other areas.

In summary, English Profile has begun 'a systematic and empirically-based approach to specifying more precisely how the CEFR can be operationalised for English, and this in turn will lead to better and more comprehensive illustrative descriptors … . In this way the CEFR will become the really useful tool that it was intended to be' (Milanovic 2009: 5).

Description or Prescription?

Can English Profile provide a definitive set of RLDs that should be learnt, taught or assessed at each level? The answer is 'no' as this is not the intention of the Programme, which seeks to describe rather than prescribe. The exact choice of language points that suit a particular course of study or other purpose will depend on a range of factors, such as:

- the level and range of levels of learners on the programme
- the age and educational background of the learners
- their reasons for learning English
- their areas of interest
- their first language
- their experience of learning English so far
- other sources of input and opportunities to practise English.

The ELT professional will need to weigh these factors up in making decisions about what to include in a course. English Profile aims to provide resources to help those decisions, including examples of curricula that have been drawn up using EP resources.

How can teachers and learners benefit from English Profile?

There is likely to be a widespread impact on English language teaching and learning in general through the support English Profile will give to those people who plan curricula, produce learning materials and design English language tests. English Profile aims to help them to judge better what language to include at each level of learning. These improvements should increase the relevance and efficiency of language learning. Individual teachers can access English Profile resources directly if they want to get more information about what is suitable for their learners at each CEFR level, see Section 5 for some ideas and examples.

We present below the main findings from research to date in the areas of Grammar, Functions and Vocabulary.

2 The English Grammar Profile

As mentioned above in the section on the development of the CEFR, the CEFR distinguishes six levels of proficiency in terms of functions, that is the different uses to which language can be put and how learners learn to use language in different ways as they master a second (or subsequent) language. The CEFR does this by providing illustrative descriptors in the form of *Can Do statements* (given in Table 1 above), for example, A2 level learners 'Can understand sentences and frequently used expressions related to areas of most immediate relevance' whereas at the next level up, B1, learners 'Can understand the main points of clear standard input on familiar matters regularly encountered in work, school, leisure, etc'.

These illustrative descriptors do not provide any language-specific details about the grammar and lexis that are characteristic of each proficiency level as the CEFR's authors intended it to be neutral with respect to the L2 being acquired so as to be widely compatible across European languages, for example allowing a given level of proficiency in L2 German to be comparable with a corresponding level in L2 French or English. The CEFR is therefore deliberately "underspecified" with respect to key linguistic features that teachers or assessors look for in a learner's L2 production when they seek to assign a particular proficiency level (see Milanovic 2009). Learners can perform any of the functions in the CEFR's illustrative descriptors by using a wide variety of grammatical constructions and words, and the ability to "do" the task does not indicate precisely *how a learner does it* or with *what grammatical and lexical properties* of English (or of another target language).

With respect to grammar the EP Programme aims to provide reference level descriptions (RLDs) that include grammatical and lexical details of English to extend the CEFR's functional characterisation of the different levels, primarily using the Cambridge Learner Corpus (CLC). This is done via *criterial features*, defined on page 6 and further exemplified below, which is central to the approach taken within the EP Programme to specify the reference levels for English.

The basic idea behind the criterial features concept is that there are certain linguistic properties that are characteristic and indicative of L2 proficiency at each level, on the basis of which teachers and examiners make their practical assessments (alongside the extent to which a learner fulfils the communicative functions required by the activity undertaken). In the area of grammar, English Profile aims to establish which grammatical features distinguish learners at each level of the CEFR, in other words what structures are used, for example, by students at B1 level but are not fully acquired by learners at A2 level and, similarly, what structures do B1 learners not appear to master until they have progressed to B2 level. The explicit identification of these distinguishing properties for grammar and lexis will add the necessary specification of linguistic features to the CEFR's functional descriptors for each of the levels. For grammar this will have considerable practical benefits for the learning, teaching and assessment of English, together with associated materials development and curriculum planning activities.

In what follows, we provide a description about what learners know in terms of grammar at A2, B1, B2, C1 and C2 levels on the basis of a list of linguistic features we have identified as being "criterial", i.e. indicative of these levels. We should note that criterial features are linguistic properties that are not meant to capture all language features that a learner uses at a certain level but distinguish a level from adjacent higher and lower levels. In the area of grammar, we focus on two types of criterial features (Hawkins & Buttery 2009, 2010; Hawkins & Filipović 2011):

(i) *correct linguistic properties* of English that have been acquired at a certain level and that generally persist at higher levels (see Tables 2-7)

(ii) *incorrect properties or errors*, that occur at a certain level or levels and with a characteristic frequency (see Tables 8-11).

Both the presence or absence of the errors, and especially their characteristic frequency, can be criterial for a level. It should also be emphasized, though, that no single feature can be criterial or distinctive for a whole level; only clusters of features have the potential to be criterial for a level.

The details of the relationships between structures and CEFR level have been identified by careful analysis of the Cambridge Learner Corpus (CLC),which has, so far, been the most important data source for English Profile Grammar researchers. The CLC is a large collection of exam scripts written by students taking Cambridge ESOL English exams around the world and has been collected since the early 1990s in a collaborative project by Cambridge University Press and Cambridge ESOL. It grows each year by around 3 million words, and currently contains:

- over 45 million words from more than 200,000 learners
- over 333,000 exam responses
- all CEFR levels
- 138 first languages
- English from learners in 203 countries.

The CLC has some important features which are not found in other L2 learner corpora, including the nature of its raw data, its size and structure, and how the material has been error-coded, tagged and parsed. The exam scripts have been systematically categorised by their CEFR level according to reliable information captured during the examination process together with a large amount of demographic information about the learners, including their L1, age and language learning motivation. This background information about the learners captured alongside their responses (and the questions they responded to) allows researchers to compare different L1 learners with respect to the English that they produce.

Half of the corpus has been painstakingly coded for errors by an expert team at Cambridge University Press and the corpus has been tagged and parsed by computational linguists at the University of Cambridge Computer Laboratory. Uniquely, over 21 million words of the CLC have been coded for errors using a coding system devised by Cambridge University Press (see Nicholls 2003). These error codes are added to the corpus manually by a team of language specialists who have been working with and refining the system for over 15 years. The CLC error-coding system is therefore one of the most comprehensive and precise systems available, which means that researchers can quickly and accurately answer questions like:

- **How often do students make a particular type of mistake?** Frequency information tells us whether a particular type of mistake is common so that we can tell which are the most problematic areas for students.

- **Which students make which mistakes?** We have a lot of information about the learners' backgrounds in English language learning, including information about their first language. This allows researchers to work out which areas of English are particularly problematic, for example, for speakers of Spanish, as opposed to Chinese or Russian speakers, so that our profile of English can accurately reflect the particular needs of learners everywhere.

In addition to being searchable lexically and by error type, the CLC is, to the best of our knowledge, the only learner corpus of English that can also be searched syntactically, i.e. for specific grammatical relations and patterns. This is possible because the CLC has been syntactically analysed using RASP (Robust Accurate Statistical Parser), a state-of-the-art automatic text parsing system developed at the University of Cambridge Computer Lab (Briscoe, Carroll & Watson 2006) and further advanced within the context of English Profile Programme (see Hawkins & Buttery 2009). Combining grammatical information and statistical patterns, the RASP system assigns part-of-speech tags to individual words, groups words into phrases by creating parse trees and extracts the grammatical relations between these phrases. The application of RASP to the CLC by a team of computational linguists enables detailed and highly accurate syntactic analyses of learner English, allowing a mapping of learner syntax and error patterns across the CEFR levels.

The findings reported here are mostly from Hawkins & Filipović (2011) but also from O'Keeffe & Mark (in preparation), defined in pedagogic terms and enriched with CLC learner examples by the Cambridge English Profile

team. These findings reflect the CLC as it currently stands. Since English Profile is a long-term programme of research, the list of criterial features presented here will be enhanced and amended as more examination and non-examination learner corpus data become available to the Programme (see Section 6 for further details).

Grammatical criterial features

This section provides a comprehensive inventory of all grammatical criterial features identified thus far and gives four types of information:

a) a summary of grammatical criterial features that distinguish each level, with example utterances from learner English (Table 2), followed by a more comprehensive description at each level

b) examples of more detailed grammatical Reference Level Descriptors (Tables 3-7)

c) examples of error types that significantly improve between adjacent levels (Tables 8-11)

d) an example of how a grammatical feature develops in learner language across the CEFR levels (O'Keeffe & Mark in preparation) (Table 12).

a. Grammatical criterial features that distinguish each level

Table 2: Key distinguishing features of learner English by CEFR level

Level	Some key features	Examples from the CLC at the appropriate level
A2	- Simple sentences - Sentences with clauses joined by *that* - Descriptive phrases introduced by a past participle - Simple direct *wh-* questions - Simple sentences using infinitives - Other infinitives - Some modals	*We came back and went to bed* *I know that you have a new house too* *There are beautiful paintings painted by famous Iranian painters* *What are you going to wear?* *I want to buy a coat* *… something to eat* *We must be there at 7 o'clock in the morning.*
B1	*-ing* clauses *Whose* relative clauses - Indirect questions - Clauses with *what* as subject/object - Verb+object+infinitive *easy* + infinitive - Some complex auxiliaries - Additional modal uses	*Maria saw him taking a taxi* *… this famous painter whose pictures I like so much* *Guess where it is?* *This is what I think* *I ordered him to gather my men.* *The train station is easy to find.* *would rather, had better* *I have invited all his friends, so we should be 28 people.*
B2	-ing clause before the main clause *It* + verb + infinitive phrase *Wh*-clause as subject of main clause - Reported speech - Lexically-specific verbs/adjectives + infinitive	*Talking about spare time, I think we could go to the Art Museum* *It would be helpful to work in your group as well.* *What came after was what really changed my summer!* *I told him I loved his songs.* *… proved to be wrong, turned out to be …, expected to …*

C1	• Lexically-specific verbs + object + infinitive • *Might* for permission • Fewer grammatical errors with agreement, countability or word formation	• *I believe her to be this country's best representative.* • *Might I tell you what we [should/will] discuss?*
C2	• Some new lexically-specific verbs + object + infinitive • Longer utterances with greater accuracy	• *They declare some products to be the hits of the season*

So what language features could be said to criterial, i.e. distinguish a learner at each level? The features listed below are most – but not all – of the structural features that the corpus-informed research has shown to be significant for each level, from A2-C2 level (for a fuller inventory of grammatical criterial features identified thus far see Tables 3-7 below). The focus is on grammar rather than vocabulary although the use of specific lexical items is also mentioned. We do not provide criterial features for A1 level, as investigation of this level is underway; our initial research suggests that some basic structures are already attested at this level. Currently, the most complete description of the A1 level for English is provided by Breakthrough (Trim 2009), available on the EP website.

Structural features that are significant for each level (from A2-C2 levels)

> **A2**
>
> By the time the learner reaches A2, certain language features are emerging as being criterial, distinct characteristics. The average length of utterance is 7.9 words (based on learner data from the Cambridge Learner Corpus). These consist usually of simple sentences such as *We came back and went to bed* (see examples in A2.1 in the table below), *I met a lot of interesting people* (A2.2), or *I can give you my guitar* (A2.3). More complex sentences produced by A2 learners are of the type *I knew that you have a new house* or *I think the zoo is an interesting place*. As these examples show, the *that* which may be used to join clauses in sentences of this type may be included or omitted by learners at this level (A2.4). Users at this level are also able to write sentences such as *There are beautiful paintings painted by famous Iranian painters* where there is a descriptive phrase introduced by a past participle (A2.8). However, similar descriptive phrases introduced by a present participle are a characteristic of a B1 rather than an A2 user (B1.3).
>
> A2 learners are able to produce simple direct *wh-* questions as well as statements (A2.6). For example, *What are you going to wear? How did you know I liked skateboards?* They are also able to produce simple sentences using infinitives – *I want to buy a coat* or *I would like to sell a book*, for example (A2.5). Infinitives are also found to be used correctly by A2 users in phrases like *something to eat* (A2.7).
>
> A2 users also have gained the ability to use some modals in some of their basic senses. They can for instance use *may*, *can* and *migh*t for possibility (A2.10), *must* for obligation (A2.11) and *should* for advice (A2.12).
>
> In other words an A2 level of English is characterised by the use of basic and relatively simple structures.

B1

As far as B1 is concerned, it can first of all be noted that learners typically continue to write steadily longer sentences as they go up the levels. The mean length of a B1 utterance is 10.8 words in comparison with 7.9 at A2 level and 14.2 at B2 level. One reason for the increasing length is that learners are able to cope with more complex sentences. For example, they produce sentences containing -*ing* clauses, which follow and complement the direct object of a main clause: *Maria saw him taking a taxi* (B1.2) or which function as adverbs and follow the main clause they modify: *He was sitting there, drinking a coffee and writing something* (B1.11). Interestingly, placing the –*ing* clause before the main clause as in *Talking about spare time, I think we could go to the Art Museum*, is a feature that characterises B2 users of the language (B2.1).

There are a number of other types of complex sentence that appear to be criterial at B1 level. While learners may cope well with simple relative clauses using *who* and *which* at lower levels, the use of *whose* as a relative pronoun is typical of this level e.g. *I met a very nice boy whose name's John* (B1.6). Indirect questions similarly appear to be a marker of B1 language e.g. *Guess where it is, I don't know what to do* (B1.8-9). Simple clauses beginning with *what* (…*what I think, What I saw*…) acting as either the object or subject of a main sentence are also successfully used by learners at this level, for example, *This is what I think, What I saw was so amazing* (B1.7).

B1 learners are getting more confident in using the infinitive in more complex structures. While A2 level students cope well with *verb + infinitive* (e.g. *I would like to buy…* or *I want to sell…* A2.5)*,* at B1 level students make use of structures consisting of verb + object + infinitive, with or without *to* e.g. *I would like you to come, I want you to do it, I helped him bake the cake* (B1.1). In addition, there are a number of other uses of infinitives linked with specific vocabulary items that seem to feature as typical features of a learner having reached B1 (B1.12-14). The level-significant word in each case is bolded in the following examples: *Monica **seems** to be good. I was **supposed** to go to the English class* (B1.12). *Your friends **expect** you to spend an amazing holiday with them. I would **like** you to spend a weekend at my house. I **want** you to say hi [to] everybody* (B1.13). *The train station is **easy** to find* (B1.14)*.*

Different auxiliary and modal uses help to map a learner's progress from one level to another. As far as B1 is concerned, the complex auxiliaries *would rather* and *had better* are typical features of a B1 learner (B1.10). So also are certain modal uses such as *may* for permission (*May I borrow* ….), *must* for logical necessity (*He is having a great time and must be really happy there*) and *should* for probability (*I have invited all his friends, so we should be 28 people*) (B1.16-18). The number of error types that significantly improve from A2 to B1 is fairly modest, and improvements focus on the use of quantifiers (e.g. *some, a lot of, any, several, other, every* etc.).

To sum up, the B1 language user has mastered the basic structures of the language and is beginning to attempt to produce more complex language.

B2

The mean length of a B2 user's utterance is 14.2 words, an increase of 3.4 words on B1. This can be explained in part by the fact that increasingly complex sentences are used as the learner progresses through the levels.

B2 language use typically, for example, may include sentences beginning with an adverbial clause introduced by an *-ing* word (B2.1) e.g. *Talking about spare time, I could go to the Art Museum*.

Another example of a more complex structure which characterises a B2 user is a sentence introduced by *It* and followed by an infinitive phrase e.g. *It would be helpful to work in your group as well* (B2.2). Yet another B2 structure is a sentence beginning with a *Wh-* word introducing a clause acting as the subject of the main clause of the sentence (B2.3) e.g. *What attracted me the most was the possibility of meeting people of the same interests.* A third example of a structure characterising a B2 level user is a main sentence with a direct object followed by a subordinate complement clause with or without *that* (B2.4), for example, *I told him (that) I loved his songs*.

As at B1 level, there are certain verbs and adjectives which, when associated with infinitive structures, can be seen as characteristic of the level (B2.6-9); examples of some of these are given below with the significant words highlighted in bold:
*My worries **proved** to be wrong.*
*Unfortunately for me the situation **turned out** to be opposite to what I thought it was.*
*And whenever money is involved, some problems are **likely** to happen.*
*You are **sure** to arrive at work on time.*
*I would **prefer** my accommodation to be in log cabins.*
*How many hours a day should I be **expected** to work?*
*Your theatre is **known** to present excellent spectacles.*
*So zoos could be the only place where people could spend their time avoiding the pollution we are **obliged** to live with every day.*
*Woods is **thought** to stand for all of white people and this book could have an influence on them.*
*The grammar and vocabulary are a bit **hard** to learn.*

There are no instances of modal use which can be seen as being criterial for a B2 level of language. Like in B1, the number of error types that appear to improve substantially at this level remains small.

C1

At C1 level the average length of utterance continues to increase, rising from B2's 14.2 to 17.3 words.

It is noticeable once learners move into C levels that the criterial, distinguishing features of their language are a combination of lexis and structure. Learners use structures mastered at earlier levels but with a much wider range of vocabulary and in more accurate ways. Verbs such as *chance, believe, find, suppose, take, assume, discover, feel* and *prove* used in complex structures which include an infinitival clause are characteristic of the language of a C1 level learner (C1.1-3). Here are some examples (see Table 6 for more):

- *I chanced to know about your competition from an international magazine.*
- *Being born and raised in Mexico, I believe her to be the country's best representative in the world.*
- *I can assure you that the strike isn't as worrying as you suppose it to be.*
- *Secondly, the low cost of membership and entry was assumed to be an advantage as well.*
- *The internet is a valuable tool, which can be proved to be the most important aspect in the learning process.*

The only strictly grammatical features which might be seen as criterial, new features of language as the learner moves from B2 to C1, are the use of a double *'s* genitive structure (e.g. *the bride's family's house*) (C1.4) and the use of *might* to convey the idea of permission (*Might I tell you what we discussed*?) (C1.5).

In general, however, it is not mainly the case that the C1 learner is mastering new grammatical features, it is more that he or she is using those that were already available at B2 level in a much more accurate way. Far fewer grammatical errors with, say, agreement, countability or word form occur at C1 than at B2 level.

C2

The trends noted at C1 continue into C2 in that what marks out C2 users of the language is the fact that they have a greater grammatical accuracy and a wider lexical range than C1 users; in other words there are no specific new structures appearing at this level. C2 users also continue the trend of producing a longer average utterance than at lower levels; the mean length of utterance for C2 is 19 words as opposed to 17.3 at C1.

Lexical range in relation to specific structures continues to expand. For example *declare, presume, remember* and the adjective *tough* are used with the following complex infinitive complements at C2 level (C1.1-3):

They declare some products to be the hits of the season, thus creating fashion and few of us want to be unfashionable.

- *He presumed work to be the way to live.*
- *Not only meetings with new people are presumed to give new experiences.*

She remembered her father to be a lively, tall and broad shouldered man with a beard that tickled when he bent down to kiss her goodnight.

- *What she knew would be really tough to live with was the reason of his death.*

At the same time, the C2 level sees the highest number of significant grammatical error improvements of all CEFR levels. Like in C1, there are significant reductions of error rates for most of the error types identified in the Cambridge Learner Corpus. Overall, it can be said that it is at the C levels where learners appear to be mastering and accurately using the majority of grammatical features in English.

To sum up, as learners progress from level A1 through to B2, they gradually acquire new structures which can be identified as characteristic of each level. Once they reach C levels, learners' progress is characterised by increased structural accuracy and by greater lexical accuracy and range rather than by the addition of new structures to their repertoire.

b. Grammatical Reference Level Descriptors for English

Table 3: Grammatical criterial features for A2 level

	Structure/feature	Examples from the CLC (A2 level)
A2.1	**Simple clauses without object** **[Intransitive clauses, NP-V]** These are simple clauses consisting of a Noun Phrase and an intransitive Verb, i.e. a verb that takes no object (e.g. *go, arrive*). The Noun Phrase is the subject of the intransitive Verb. The intransitive Verb can be followed by a particle (e.g. *We came back*), an adverb (e.g. *You can get there*) or a prepositional phrase (e.g. *You can go to Yite Park*).	• *Can you come to my house at 2pm on Saturday?* • *You can get there by train.*
A2.2	**Clauses with a direct object** **[Transitive clauses, NP-V-NP]** These are clauses containing a Noun Phrase followed by a Verb and another Noun Phrase (e.g. *he loved her*). The Noun Phrase preceding the Verb is the subject and the Noun Phrase following the Verb is the direct object (i.e. *he* is the subject and *her* is the direct object of the verb *loved*). The Noun Phrase serving as the object of the Verb can be followed by a prepositional phrase (e.g. *Now I write a postcard for you*).	• *I met a lot of interesting people.* • *Now I write a postcard for you…*
A2.3	**Clauses with a direct and indirect object** **[Ditransitive clauses (i) NP-V-NP-NP & (ii) NP-V-NP-PP (P=to)]** (i) These clauses comprise a Noun Phrase, a Verb and two Noun Phrases after the Verb (e.g. *I can give you my guitar*). The Noun Phrase preceding the verb is its subject (e.g. *I*). The first Noun Phrase following the Verb is the indirect object (e.g. *you*) and the second Noun Phrase the direct object (e.g. *my guitar*). (ii) In these clauses (*I'll give it to you*) the Verb is followed by a Noun Phrase which is the direct object (e.g. *it*) and a prepositional phrase with *to* which is *the indirect object (e.g. to you)*.	• *I can give you my guitar.* • *I'll give it to you…*
A2.4	**Verb + subordinate clause with or without *that*** **[Verb with a finite complement clause NP-V-S]** These structures comprise one main clause containing a Noun Phrase and a Verb and a subordinate complement clause with a finite Verb (i.e. a verb inflected for person and tense) (e.g. *they thought that he was always late*). *that* at the beginning of the subordinate clause can be overt or can be omitted (e.g. *I hope (that) you are well*).	• *I knew that you have a new house too.* • *I think the zoo is an interesting place.*
A2.5	**Verb + infinitive** **[Verb with subject-controlled infinitival complements, NP-V-VPinfin]** These are clauses comprising a Noun Phrase and a main Verb, which is followed by another Verb in the infinitive (to+Verb) (e.g. *I wanted to come*). The infinitival Verb is the complement of the main Verb. The Noun Phrase is the subject of both the main Verb and of the infinitival Verb.	• *I want to buy a coat.* • *I would like to sell a book.*
A2.6	**Direct WH-questions** These are questions beginning with a WH-word (e.g. *what, who, where*) followed by an auxiliary (e.g. *have, do*) or a copula (e.g. *be*), a Noun Phrase (e.g. *you, the park*), and a Verb (e.g. *What do you like?*).	• *What are you going to wear?* • *So when are you celebrating your birthday?*
A2.7	**Pronoun + infinitive** These are phrases beginning with a Pronoun (e.g. *something*) followed by a Verb in the infinitive (*to+*Verb) (e.g. *something*	• *You can bring something to eat if you*

		Examples
	to drink).	• *want to.* • *... I'm going to a farm this weekend and there's **nothing to do**.*
A2.8	**Noun + descriptive phrase introduced by past participle [Postnominal modification with –ed]** These are phrases comprising a Noun Phrase followed by a Past Participle ending in –ed (e.g. ***the boy deprived** of his ice cream*).	• *There are beautiful **paintings painted** by famous Iranian painters*
A2.9	**Double embedded genitive with (of... (of...))** **[An [of [of]] double embedding of postnominal genitives]** These are complex phrases containing a Noun Phrase (e.g. *the beginning*) followed by an –of genitive phrase (e.g. *of the end*), which contains another –of genitive phrase (e.g. *the beginning **of the end of civilization***).	• *I like the colours **of the back of the mobile phone**...*
A2.10	**Modals MAY, CAN and MIGHT in the Possibility (epistemic) sense** The modal verbs *may, can* and *might* are used to denote Possibility (e.g. *It **may** rain this afternoon*).	• *Then we **may** go sightseeing.* • *... the paint **might** make our t-shirts dirty.*
A2.11	**Modal MUST in the Obligation (deontic) sense** The modal verb *must* is used to denote Obligation (e.g. *You **must** not drive so fast*).	• *We **must** be there at 7 o'clock in the morning.*
A2.12	**Modal SHOULD in the Advice (deontic) sense** The modal verb *should* is used to denote Advice (e.g. *You **should** see a doctor*).	• *You **should** wear old clothes because we will get dirty.*

Table 4: Grammatical criterial features for B1 level

	Structure/feature	Examples from the CLC (B1 level)
B1.1	**Verb + object + infinitive** **[Verbs with object-controlled infinitival complements NP-V-NP-VPinfin]** These are complex clauses comprising a main clause with a Verb followed by a Noun Phrase and a complement clause with an infinitival Verb (to+Verb) with or without to (e.g. **I helped her (to) bake the cake**). The Noun Phrase her is the direct object of the main verb and the logical subject of the infinitival Verb.	• **I called my assistant and ordered him to gather** my men to the hall.
B1.2	**Verb + object + Verb ending in –ing** **[Verbs with object-controlled –ing complements NP-V-NP-VPing]** These are clauses beginning with a Noun Phrase and a Verb followed by another Noun Phrase and a Verb ending in –ing (e.g. **I caught him stealing**). The Noun Phrase him is the direct object of the main Verb and the logical subject of the Verb ending in –ing.	• **Maria saw him taking** a taxi.
B1.3	**Noun + descriptive phrase introduced by present participle** **[Postnominal modification with –ing]** These are phrases consisting of a Noun Phrase followed by a Present Participle ending in -ing (e.g. **the boy walking** down the road).	• As I was desperate I put an **advertisement in the newspaper asking** if someone had it, but no one answered me.
B1.4	**It + Verb + subordinate clause with or without that** **[It Extraposition with finite clauses]** These are complex clauses containing one main clause beginning with It and a Verb Phrase, followed by a subordinate complement clause starting with that and containing a finite Verb (i.e. a verb that is inflected for person or tense) (e.g. **It doesn't matter that she is not here**). The use of that is optional (e.g. **It is a pity (that) they cannot come**).	• **It's true that I don't need** a ring to make me remember you.
B1.5	**Verb + Prepositional Phrase + subordinate clause with or without that** **[Verbs with a PP plus finite complement clause, NP-V-PP-S]** These are complex clauses beginning with a Noun Phrase followed by a main Verb that takes as complements a Prepositional Phrase and a subordinate complement clause starting with that and containing a finite Verb (i.e. a verb that is inflected for person or tense) (e.g. **They admitted to the authorities that they had entered illegally**). The Prepositional Phrase to me is the indirect object of the main Verb admitted and the subordinate clause is the direct object of the Verb. that at the beginning of the subordinate clause can be overt or can be omitted (e.g. **They admitted to the authorities (that) they had entered illegally**).	• **He said to me he would like** to come back soon.
B1.6	**Relative clauses with whose** **[Relative clauses formed on a genitive position]** These are relative clauses beginning with whose (i.e. WH-word in the genitive) and are followed by another Noun Phrase (e.g. **whose pictures**).	• … a biography of this famous painter **whose pictures** I like so much
B1.7	**WH-word + Noun Phrase + Verb clauses used as subject or object** **[Pseudoclefts type (i) WH-NP-VP]** These are subordinate clauses beginning with a WH-word (e.g. what) followed by a Noun Phrase and a Verb (e.g. what I like). The WH-word is the direct object of the Verb in this clause (i.e. what is the object of like). The Pseudocleft type (i) structure typically appears within another sentence as a subject (**What I like is watching football**) or as an object (complement) after the Verb be (This is **what I like**)	• **What I absolutely dislike** is go shopping, although I really like new clothes. • I opened the door and **what I saw** was so amazing.

		Examples
B1.8	**Indirect WH-questions** These are subordinate clauses beginning with a WH-word (e.g. *what*) followed by a Noun Phrase and a finite Verb, i.e. a verb inflected for person and tense (e.g. *he asked what he should do*).	• *Guess **where it is**.* • *I don't know **how I could have done it**.*
B1.9	**Indirect WH questions with infinitive** These are subordinate clauses beginning with a WH-word (e.g. *what*) followed by a Noun Phrase and an infinitival Verb (to+Verb) (e.g. *he explained how to do it*).	• *I did not know **where to look** for it anymore.* • *I don't know **what to do**.*
B1.10	**Complex auxiliaries WOULD RATHER and HAD BETTER** These are complex auxiliaries that consist of a Simple Auxiliary and an Adverb (e.g. *would rather, had better*) and are followed by a Verb (e.g. *I **had better improve** my game*).	• *They won't be very happy if you go on holiday with them and then you are bad tempered all the time because **you would rather be** with your friends.* • *If you don't like to go with them **you had better tell** them why you don't want to come.*
B1.11	**Adverbial subordinate clauses with –*ing* that follow the clause to which they are attached** These are subordinate clauses with verbs ending in –*ing* (e.g. *walking down the road*). These subordinate clauses have an adverbial function and can follow the main clause (e.g. *he pumped into a lamppost **walking down the road***).	• *He was sitting there, **drinking** a coffee and **writing** something.*
B1.12	***seem, supposed* (adjective) + infinitive** **[Subject-to-Subject Raising, NP-V-VPinfin]** These are complex clauses containing a main clause beginning with a Noun Phrase and followed by the Verb *seem* (e.g. *John seems*) or the Verb *be* and the Adjective *supposed* (e.g. *John is supposed*) and a subordinate clause with an infinitival Verb (to+Verb) (e.g. ***John seems/is supposed to be** good*). In the Subject-to-Subject Raising constructions, the subject of the main Verb is also the logical subject of the infinitival Verb and it has moved out of its subject position within the subordinate clause to become the subject of the main clause (hence 'Subject-to-Subject Raising').	• ***Monika seems to be** good, intelligent teacher, but I have a bad feeling about Paula.* • *It looks like I have a rehearsal for the "Yamaha-Concert" at the same time **I was supposed to go** to the English class.*
B1.13	***expect, like, want* + object + infinitive** **[Subject-to-Object Raising, NP-V-NP-VPinfin]** These are complex clauses containing a main clause beginning with a Noun Phrase and the Verbs *expect, like* or *want* (e.g. *John expects*), followed by a Noun Phrase and a subordinate clause with an infinitival Verb (to+Verb) (e.g. ***John expects her to come** to the party*). The Noun Phrase *her* is the direct object of the main Verb and the logical subject of the subordinate infinitival clause, and it has moved out of its position as the subject to become the direct object of the main clause (hence 'Subject-to-Object Raising').	• *Sara told me she would come, but **I didn't expect her to come** so early.* • *I am doing fine, **I expected it to be** more difficult, but it is not so hard.* • *Finally **I want you to say** hi [to] everybody and I am looking forward to seeing you.*

B1.14	*easy + infinitive* **[Tough Movement with the adjective *easy*]** These are complex clauses comprising a main clause beginning with a Noun Phrase, the Verb *be* and the adjective easy (e.g. *the book is easy*) followed by a subordinate clause with an infinitival Verb (to+Verb) (e.g. ***The book is easy to read***). In Tough Movement structures the subject of the main clause is the logical object of the subordinate infinitival clause, which has moved from the object position (e.g. *to read the book*) to the position of the subject of the main clause. The book is the logical object of reading and this whole event (reading the book) is claimed to be easy, not necessarily the book itself (e.g. *the book could be easy to read but hard to carry*). The raised Noun Phrase in Tough Movement constructions can occupy other non-subject positions in the subordinate clause e.g. *The student is easy to study with*	• *The train station is **easy to find**.* • *The problem you have is not very **easy to solve**.*
B1.15	**Double embedded genitive with (*of... (-'s)*)** **(An [*of* [-s]] double embedding of an *–s* genitive within an *of* genitive)** These are complex phrases containing a Noun Phrase (e.g. *the beginning*) followed by an *–of* genitive phrase which contains an *–s* genitive (possessive) phrase (e.g. *the beginning of **the professor's book***).	• *I am a big fan **of the world's** most famous British secret service agent.*
B1.16	**Modal MAY in the Permission (deontic) sense** The modal verb MAY is used to denote Permission.	• ***May** I borrow your bicycle for this weekend?*
B1.17	**Modal MUST in the Necessity (epistemic) sense** The modal verb must is used to denote logical or physical Necessity.	• *This movie **must** be great. My sister told me that she was amazed [when] she [saw] the film so I think it's really good.* • *He is having a great time and **must** be really happy there.*
B1.18	**Modal SHOULD in the Probability (epistemic) sense** The modal verb should is used to denote Probability.	• *I have invite[d] all his friends, [so] we **should** be 28 people.*

Table 5: Grammatical criterial features for B2 level

	Structure/feature	Examples from the CLC (B2 level)
B2.1	**Adverbial subordinate clauses with *–ing* that precede the clause to which they are attached** These are subordinate clauses with verbs ending in *–ing* (e.g. *walking down the road*). These subordinate clauses precede the main clause and act like adverbs modifying the main clause (e.g. *Walking down the road, he pumped into a lamppost*).	• *Talking about spare time, I think we could go to the Art Museum*
B2.2	***It* + Verb + infinitive** **[*It* Extraposition with infinitival phrases]** These are complex clauses containing one main clause beginning with *It* and a Verb, followed by a subordinate complement clause with a Verb in the infinitive (to+Verb) (e.g. *It is likely to rain tomorrow*).	• *…it would be helpful to work in your group as well.*
B2.3	***WH-word + Verb Phrase* clauses used as subject or object** **[Pseudo-cleft type (ii) WH-VP]** These are subordinate clauses beginning with a WH-word (e.g. *what*) followed by a Verb and often a Noun Phrase (e.g. *what interests me*). The WH-word is the subject of the Verb in this clause (i.e. *what is the subject of interest*). The Pseudocleft type (ii) structure typically appears within another sentence as a subject (*What interests me is politics*) or as an object (complement) (*I know what interests me*).	• *What attracted me the most was the possibility of meeting people of the same interests.* • *But what came after was what really changed my summer!*
B2.4	**Verb + object + subordinate clause with or without *that*** **[Verbs with an NP plus finite complement clause, NP-V-NP-S]** These are complex clauses comprising one main clause beginning with a Noun Phrase and a Verb followed by a Noun Phrase and a subordinate complement clause (e.g. *he told me that the audience was leaving*). The subordinate complement clause can be optionally introduced by *that*.	• *I told him I loved his songs* • *She told me that she had worked for summer camp for children*
B2.5	**Verb + object + adjective** **[Secondary predications, NP-V-NP-AdjP]** These are clauses comprising a Noun and a Verb, which is followed by a Noun Phrase and an Adjective (e.g. *he painted the car red*). The Noun Phrase following the Verb is the direct object (*the car*) and the adjective (*red*) complements that object. In Secondary Predication constructions, the direct object of the verb (*the car*) also contracts a secondary relation with the following predicate (*is or becomes red*).	• *But if you don't want to take any risks, just go and paint the houses yellow and blue*
B2.6	**the verbs *appear, cease, fail, happen, prove, turn out,* and the adjectives *certain, likely, sure, unlikely* + infinitive** **[Subject-to-Subject Raising, NP-V-VPinfin]** These are complex clauses containing a main clause beginning with a Noun Phrase and followed by the Verbs *appear, cease, fail, happen, prove, turn out* (e.g. *The noise ceased*) or the Verb *be* and the Adjectives *certain, likely, sure, unlikely* (e.g. *John is likely*) and a subordinate clause with an infinitival verb (to+Verb) (e.g. *The noise ceased to get on his nerves*). In the Subject-to-Subject Raising constructions, the subject of the main Verb is also the logical subject of the infinitival Verb and it has moved out of its subject position within the subordinate clause to become the subject of the main clause ('Subject-to-Subject Raising').	• *To my regret, the evening totally failed to live up to my expectations.*

B2.7	***imagine, prefer* + object + infinitive** **[Subject-to-Object Raising, NP-V-V-NP-VPinfin]** These are complex clauses containing a main clause beginning with a Noun Phrase and the Verbs *imagine* or *prefer* (e.g. *I would prefer*), followed by a Noun Phrase and a subordinate clause with an infinitival Verb (to+Verb) (e.g. *I would prefer my accommodation to be* in log cabins). The Noun Phrase *my accommodation* is the direct object of the main Verb and the logical subject of the subordinate infinitival clause, and it has moved out of its position as the subject of the subordinate clause to become the direct object of the main clause ('Subject-to-Object Raising').	• *You were so kind and friendly that I **had never imagined myself to have visited** far foreign country, Alaska.* • *I would prefer my accommodation to be in log cabins, because I am allergic to some insects that might go in the tent.*
B2.8	**the verbs *expected, known, obliged, thought* (in Passive voice) + infinitive** **[Subject-to-Object Raising plus Passive, NP-V-VPinfin]** These are complex clauses containing a main clause beginning with a Noun Phrase and the Verbs *expect*, *know*, *oblige* or *think* in the passive form (e.g. ***Smoking is known***), followed by a subordinate clause with an infinitival Verb (to+Verb) (e.g. ***Smoking is known to cause** cancer*). The Noun Phrase *Smoking* is the subject of the subordinate clause that has been raised into the higher object position (e.g. *smoking* as in *We know smoking to cause cancer*) (hence 'Subject-to-Object Raising'). It is then further promoted to subject position in the main clause by the operation of the Passive.	• ***Your theatre is known to present*** *excell...* *spectacles.* • *So zoos could be the only place where people could spend their time* • *avoiding the pollution **we are obliged t... live** with every day.*
B2.9	***difficult, good, hard* + infinitive** **[Tough Movement constructions with the adjectives *difficult, good, hard*]** These are complex clauses comprising a main clause beginning with a Noun Phrase, the Verb *be* and the adjectives *difficult*, *good* or *hard* (e.g. *Apples are good*) followed by a subordinate infinitival clause (e.g. ***Apples are good to eat***). In Tough Movement structures the subject of the main clause is the logical object of the subordinate infinitival clause, which has moved from the object position (e.g. *to eat apples*) to the position of the subject of the main clause. *Apples* is the logical object of eating and this whole event (*eating apples*) is claimed to be *good*, not necessarily the *apples* themselves. The raised Noun Phrase in Tough Movement constructions can occupy other non-subject positions in the subordinate clause e.g. *The student is easy to study with*.	• *The first Restaurant is "Mexico", it is situated near the metro station and is **not very difficult to find**.* • ***The grammar and vocabulary are a bit hard to learn**.*
B2.10	**Double embedded genitive with ((of...) –'s)** **(An [[of] -s] double embedding of an *of* genitive within an -s genitive)** These are complex phrases containing a Noun Phrase (e.g. *war*) preceded by an -s genitive (possessive) phrase which contains an –*of* genitive phrase (e.g. ***the king of England's war***).	• *After that I went to **a friend of mine's house** where I spent one week.*

Table 6: Grammatical criterial features for C1 level

	Structure/feature	Examples from the CLC (C1 level)
C1.1	**the verb *chance* + infinitive** **[Subject-to-Subject Raising, NP-V-VPinfin]** These are complex clauses containing a main clause beginning with a Noun Phrase and the Verb *chance* (e.g. *I chance*), followed by a subordinate clause with an infinitival Verb (to+Verb) (e.g. *I chanced to know about your Competition from an international magazine*). In the Subject-to-Subject Raising constructions, the subject of the main Verb is also the logical subject of the infinitival Verb and it has moved out of its position as a subject of the subordinate clause to become the subject of the main clause (hence 'Subject-to-Subject Raising').	• *Dear Sirs, I chanced to know about your Competition from an international magazine.*
C1.2	**believe, find, suppose, take + object + infinitive** **[Subject-to-Object Raising, NP-V-NP-VPinfin]** These are complex clauses containing a main clause beginning with a Noun Phrase and the Verbs *believe, find, suppose* or *take* followed by a Noun Phrase and a subordinate clause with an infinitival Verb (to+Verb) (e.g. *I find this to be* more interesting than the walking route to Lake Hawksmere). The Noun Phrase *this* is the direct object of the main Verb *find* and the logical subject of the subordinate infinitival clause, and it has moved out of its position as the subject of the subordinate clause to become the direct object of the main clause ('Subject-to-Object Raising').	• *I find this to be more interesting than the walking route to Lake Hawksmere.* • *Overall I found this to be pretty satisfying as it does fulfill most of the students' wishes.*
C1.3	**assumed, discovered, felt, found, proved (in Passive voice) + infinitive** **[Subject-to-Object Raising plus Passive, NP-V-VPinfin]** These are complex clauses containing a main clause beginning with a Noun Phrase and the Verbs *assumed, discovered, felt* or *found* in the passive form, (e.g. *The children stories were felt*), followed by a subordinate clause with an infinitival Verb (to+Verb) (e.g. *The children stories were felt to be* the best idea for kids). The Noun Phrase *The children stories* is the subject of the subordinate clause that has been raised into the higher object position (e.g. *We felt the children stories were the best idea for kids*) ('Subject-to-Object Raising'). It is then further promoted to subject position in the main clause by the operation of the Passive.	• *The children stories were felt to be the best idea for kids, after of course the pony rides.*
C1.4	**Double embedded genitive with ((-'s) -'s)** **(An [[-s] -s] double embedding of an –s genitive within an –s genitive)** These are complex phrases consisting of a Noun Phrase (e.g. house) preceded by an -s genitive (possessive) phrase which contains another -s genitive (possessive) phrase (e.g. *the bride's family's house*).	• *After spending the first day of their marriage in the bride's family's house*
C1.5	**Modal MIGHT in the Permission (deontic) sense** The modal verb *might* is used to denote Permission.	• *Might I tell you what we discuss?*

Table 7: Grammatical criterial features for C2 level

	Structure/feature	Examples from the CLC (C2 level)
C2.1	***declare, presume, remember + object + infinitive*** **[Subject-to-Object Raising, NP-V-NP-VPinfin]** These are complex clauses containing a main clause beginning with a Noun Phrase and the Verbs *declare, presume* or *remember* followed by a Noun Phrase and a subordinate clause with an infinitival Verb (to+Verb) (e.g. ***He presumed work to be*** *the way to live.*). The Noun Phrase *work* is the direct object of the main Verb *presumed* and the logical subject of the subordinate infinitival clause *to be*, and it has moved out of its position as the subject of the subordinate clause to become the direct object of the main clause ('Subject-to-Object Raising').	• *They declare* ***some products to be*** *the hits of the season, thus creating fashion, and few of us want to be unfashionable* • *Plus,* ***I remember my classes to be*** *very participative, and dynamic.*
C2.2	**the verb *presumed* (in Passive Voice) + infinitive** **[Subject-to-Object Raising plus Passive, NP-V-VPinfin]** These are complex clauses containing a main clause beginning with a Noun Phrase and the Verb *presumed* in the passive form, (e.g. *Not only* ***meetings with people are presumed*** *to give new experiences*), followed by a subordinate clause with an infinitival Verb (to+Verb) (e.g. *Not only* ***meetings with people are presumed to give*** *new experiences*). The Noun Phrase *meetings* is the subject of the subordinate clause that has been raised into the higher object position (e.g. *We presume meetings to be*) ('Subject-to-Object Raising'). It is then further promoted to subject position in the main clause by the operation of the Passive.	• *Not only* ***meetings with people are presumed to give*** *new experiences.*
C2.3	***tough* + infinitive** **[Tough Movement constructions with the adjective *tough*]** These are complex clauses beginning with a Noun Phrase, the Verb *be* and the Adjective *tough* (e.g. *our team was tough*) followed by an infinitival Verb (to+Verb) (e.g. ***our team was tough to beat***). In Tough Movement structures, the subject of the main clause is the logical object of the subordinate infinitival clause, which has moved from the object position (e.g. *to beat our team*) to the position of the subject of the main clause. *Our team* is the logical object of *beat* and this whole event (*beating our team*) is claimed to be *tough*, not necessarily *our team* itself. The raised Noun Phrase in Tough Movement constructions can occupy other non-subject positions in the subordinate clause e.g. *What she knew* ***would be really tough to live with*** *was the reason of his death*	• *What she knew* ***would be really tough to live with*** *was the reason of his death.*

c) Examples of error types that significantly improve between adjacent levels

Table 8: Error types that improve significantly between A2 to B1 levels

Key to four error types tables 8-11: bold = error, (word) = correction, [word] = error has been corrected, [ ] = unnecessary word removed

	Error type	Examples from the CLC (A2 level)
1.	**Anaphor Agreement** When the anaphor word is correct and the form of the anaphor is valid but wrong in the context because it does not agree grammatically with its coordinates, it is an Anaphor Agreement error.	It's three years old and **he** works very well, but I would like a new computer. (it) I have a lot of [books]. I bought **one** years ago.(them)
2.	**Form of Determiner** When the articles 'a' and 'an' are confused.	...a month ago I bought **an** handheld. (a) This is **a** interesting place! (an)
3.	**Missing Adjective** When a sentence or construction requires an adjective for completeness and that adjective has been omitted, it is a Missing Adjective error.	Here're a **lot of kinds** of animals. (lot of different kinds) The weather is fantastic, we [all have] **really fun**... (really good fun)
4.	**Missing Adverb** When a sentence or construction requires an adverb for completeness and that adverb has been omitted, it is a Missing Adverb error.	I think we will [ ] have a good time[.] Hope to **see you** (see you soon) You can **get by** car. (get there by)
5.	**Missing Conjunction (Link Word)** When a sentence or construction requires a conjunction / link word (or words) for completeness and that word has been omitted, it is a Missing Conjunction error.	**You** [want] this please call me. (If you) I hope I go by train so I will arrive at 5 o'clock, **especially I want** to play [ ] your new videogame... (especially because I want)
6.	**Missing Quantifier** When a sentence or construction requires a quantifier for completeness and that quantifier has been omitted, it is a Missing Quantifier error.	I will **buy new** trousers and a pair of shoes. (buy some new)
7.	**Inflection of Quantifier** When the learner has created a feasible but non-valid inflected form of the quantifier.	I think that we can go for a walk, ride a bike, play football and **others** sports. (other) I bought a [pair of] trouser[s] and **somes** [T-shirts]. (some)
8.	**Replace Quantifier** When a valid quantifier word in the language has been used and it is the correct part of speech but not the correct quantifier, it is a Replace Quantifier error.	I want to sell **many** dolls. (a lot of) I will move to **other** city so I want to sell it. (another)

| 9. | Unnecessary Quantifier
When an unnecessary extra quantifier has been used in a sentence or construction in such a way that it makes the sentence or construction incorrect, it is an Unnecessary Quantifier error. | *My hobbies are **such as** singing, dancing and drawing.*
*I want to sell this table because I have a new **other** table!* |

Table 9: Error types that improve significantly between B1 to B2 levels

	Error type	Examples from the CLC (B1 level)
1.	**Derivation of Conjunction (Link Word)** Where a conjunction / link word resembles, or includes the stem of, a valid word but has been incorrectly derived, usually because it has been given an incorrect affix, it is a Derivation of Conjunction error.	*I didn't do anything, **accept** [put] a new bed in my bedroom. (except)* *Now, I'm more attentive **in** reading letters... (when)*
2.	**Derivation of Determiner** Where a determiner resembles, or includes the stem of, a valid determiner but has been incorrectly derived, usually because it has been given an incorrect affix, it is a Derivation of Determiner error.	*I'm really glad, because you want to visit **mine** country. (my)* *You asked me about the best place to spend **you** time... (your)*
3.	**Form of Determiner** When the articles 'a' and 'an' are confused.	*The film has **an** happy end... (a)* *Send me **a** e-mail... (an)*
4.	**Inflection of Determiner** When the learner has created a feasible but non-valid inflected form of the determiner, usually because of a mistaken belief that the determiner must agree in number with the noun which it precedes.	*It was really interesting to hear about all the different people and **theirs** [backgrounds]. (their)*
5.	**Inflection of Quantifier** When the learner has created a feasible but non-valid inflected form of the quantifier.	*If you have **others** questions please contact me! (other)* *I think I will go [to] the park, near the zoo, where **severals** sports are [played]. (several)*
6.	**Inflection of Verb** When the learner has made a false assumption about whether a verb is regular or irregular and inflected it accordingly. Most commonly, the error is caused by putting regular inflections on irregular verbs.	*...all the walls are white and I've **hanged** [a] few paintings on them. (hung)* *After the film we had dinner and we **chated** about the film... (chatted)*

Table 10: Error types that improve significantly between B2 to C1 levels

	Error type	Examples from the CLC (B2 level)
1.	**Agreement** When the word is correct and the form of the word is valid but wrong in the context because it does not agree grammatically with its coordinates, it is an Agreement error.	*I know that **this kind** of jobs are well paid... (these kinds)* *...you can try the clothes you choose on and be sure that **it fits** you. (they fit)*
2.	**Noun Agreement** When the noun is correct and the form of the noun is valid but wrong in the context because it does not agree grammatically with its coordinates, it is a Noun Agreement error.	*If you worked with me we would spend a good summer **holidays** together. (holiday)* *This has been the main reason to ban these **kind** of places nowadays. (kinds)*
3.	**Argument Structure** Where the structure of a sentence or phrase contravenes rules of both grammar and word order it is an argument structure error.	*But please tell me **how is the weather** in Scotland at this time. (what is the weather like)* *Therefore, I think you should **pay some money back for me.** (pay me some money back)*
4.	**Countability of Determiner** When a determiner form is used which is incorrect because of the countability of the noun to which it refers, it is a Countability of Determiner error.	*Why do you give **those** information in an advertis[e]ment? (this)* *...just to find that perfect tin of beans or **a** toothpaste. (some)*
5.	**Derivation of Anaphor** Where an anaphor resembles, or includes the stem of, a valid word but has been incorrectly derived, usually because it has been given an incorrect affix, it is a Derivation of Anaphor error.	*It was very very kind of **yours** to invite us to your home. (you)*
6.	**Derivation of Conjunction** Where a conjugation / link word resembles, or includes the stem of, a valid word but has been incorrectly derived, usually because it has been given an incorrect affix, it is a Derivation of Conjunction error.	*...you are able to take books anywhere you want to [,] **even** the book is too heavy. (unless)* *I don't think television can replace books, **as well as** I don't bel[ie]ve books can replace television. (and)*
7.	**Derivation of Determiner** Where a determiner resembles, or includes the stem of, a valid determiner but has been incorrectly derived, usually because it has been given an incorrect affix, it is a Derivation of Determiner error.	*People rely on me as I always keep **me** promises. (my)* *Clara looked at **he** watch. (her)*
8.	**Inflection of Determiner** When the learner has created a feasible but non-valid inflected form of the determiner, usually because of a mistaken belief that the determiner must agree in number with the noun which it precedes.	*...I think this experience would be useful [on] **yours** campsites. (your)* *I think animals help us with understanding not only **theirs** but also our own [behaviour]. (their)*

9.	**Form of Adjective** When a valid form of the adjective (positive, comparative, or superlative) has been used but is the wrong form in the context.	*Well, in my opinion my third choice is the **better**. (best)* *... I had the **worse** time [of] my life. (worst)*
10.	**Form of Adverb** When a valid form of the adverb (positive, comparative, or superlative form) has been used but is the wrong form in the context.	*...[the] solution [that] **better** [satisfies] our needs. (best)* *But I would like to live **nearest**, and to go there by bicycle. (nearer)*
11.	**Form of Noun** When the singular form of a noun is used where the context demands a plural, or vice versa.	*She's on **holidays** next 15th of July. (holiday)* *...most of us would never have had the chance to see some **kind** of animals. (kinds)*
12.	**Form of Quantifier** When a valid form of the quantifier (singular or plural form) has been used but is the wrong form in the context.	*...in **more** of the cases ... (in **most** cases)* *The [I]nternet makes us able to follow sports events, concerts and political speeches which are **thousand** of miles away from us. (thousands)*
13.	**Form of Verb** When either of the base, -ing and to+infinitive forms of the verb have been used where another form is required.	*If you have any more questions don't hesitate **asking** me! (to ask)* *Your idea about **find** a job before starting college is great. (finding)*
14.	**Inflection of Adjective** When the learner has created a feasible but non-valid inflected form of the adjective. This code also covers instances where learners mistakenly make adjectives agree with the nouns they modify.	*...I also know that you organise conferences and **differents** activities. (different)* *The **most easiest** way... (easiest)*
15.	**Inflection of Anaphor** When the learner has created a feasible but non-valid inflected form of the pronoun.	*This [choice] doesn't give many artists the [opportunity] to [express] **themself**. (themselves)* *...each friend of **mines** ... (mine)*
16.	**Inflection of Quantifier** When the learner has created a feasible but non-valid inflected form of the quantifier.	*...we [can] find **fews** [in] the newspaper. (a few)* *...I made **severals** phone calls... (several)*
17.	**Missing Adjective** When a sentence or construction requires an adjective for completeness and that adjective has been omitted, it is a Missing Adjective error.	*Are **they at a** discount, too? (they available at a)* *If there are many cars in a city, you waste your time [ ] **standing**. (standing still)*
18.	**Missing Conjunction (Link Word)** When a sentence or construction requires a conjunction / link word (or words) for completeness and that word has been omitted, it is a Missing Conjunction error.	*So it would be **better I** [wrote] a letter (better if I)* *That was my best day ever, the **one I shared** a meal with Paul McPartney. (one when I shared)*
19.	**Missing Determiner**	*Her neighbours were **making noise** again. (making a noise)*

#	Error type and definition	Examples
	When a sentence or construction requires a determiner for completeness and that determiner has been omitted.	Playing football is good **for health.** (for your health)
20.	**Missing Preposition** When a sentence or construction requires a preposition for completeness and that preposition has been omitted.	I look **forward hearing** from you. (forward to hearing) One thing I don't **know is** the pay. (know about is)
21.	**Missing Quantifier** When a sentence or construction requires a quantifier for completeness and that quantifier has been omitted, it is a Missing Quantifier error.	The organiser gave me a radio and I had to advise the other staff members in **case of disorder.** (case of any disorder) In the **last years**, there has been a great improvement... (last few years)
22.	**Missing Verb** When a sentence or construction requires a verb for completeness and that verb has been omitted.	After that we **will back** home. (will come back) While before it **would enough** to know ... (would have been enough)
23.	**Unnecessary Conjunction (Link Word)** When an unnecessary extra conjunction / link word has been used in a sentence or construction in such a way that it makes the sentence or construction incorrect.	... I was also working with children in my town, **because** in order to get more money . She thought only magic **that** could help her but it was impossible.
24.	**Unnecessary Preposition** When an unnecessary extra preposition has been used in a sentence or construction in such a way that it makes the sentence or construction incorrect.	I am looking forward to hearing **about** your answer. Going **for** shopping is a good thing ...
25.	**Unnecessary Quantifier** When an unnecessary extra quantifier has been used in a sentence or construction in such a way that it makes the sentence or construction incorrect, it is an Unnecessary Quantifier error.	That's why I don't agree with **some** people who think that keeping animals is cruel and unnecessary. Your friends will without **no doubt** [ ] try to help you, and don't forget it is not too late yet.
26.	**Unnecessary Verb** When an unnecessary extra verb has been used in a sentence or construction in such a way that it makes the sentence or construction incorrect.	Could you tell me about how much **do** you want to get per hour ... We all decided **to go** to have a day by the seaside ...
27.	**Negative Formation** When an attempt to construct a negative results in an invalid construction.	I **hadn't** a good time! (didn't have) I could **not hardly** believe it. (hardly)
28.	**Complex Error** Where an ungrammatical and incomprehensible string of words are written and it is impossible to correct them, it is a Complex Error.	I wish you to **spend your stay as more exciting as you can.** **Without states of minds**, she succeeded in making me feel that life was going on.

Table 11: Error types that improve significantly between C1 to C2 levels

	Error type	Examples from the CLC (C1 level)
1.	**Anaphor Agreement** When the anaphor word is correct and the form of the anaphor is valid but wrong in the context because it does not agree grammatically with its coordinates, it is an Anaphor Agreement error.	I was expecting to meet a friend of mine **which** lives in St. Andrews ... (who) These two computer games are currently **one** of the best football game[s] [ ] made in this decade. (two)
2.	**Determiner Agreement** When the determiner is correct and the form of the determiner is valid but wrong in the context because it does not agree grammatically with its coordinates, it is a Determiner Agreement error.	Despite the length of the journey there was not **any** toilet available [on] the coach. (a) Fashion was a [general] word, meaning the two types of clothes which were worn in **that** days: ... those)
3.	**Verb Agreement** When the verb is correct and the form of the verb is valid but wrong in the context because it does not agree grammatically with its coordinates, it is a Verb Agreement error.	Something which grows in popularity **are** the solar cells. (is) The world **have** changed. (has)
4.	**Argument Structure** Where the structure of a sentence or phrase contravenes rules of both grammar and word order it is an argument structure error. These errors are often calques or direct translations of L1 structures.	I **demand an apology to be** published in your newspaper... (demand that an apology be) Then in the advertisement **it is written something** about a social programme which is totally absent. (something is written)
5.	**Countability of Determiner** When a determiner form is used which is incorrect because of the countability of the noun to which it refers, it is a Countability of Determiner error.	I hope you will be able to improve the programme with **these** [information]... (this) ...we are quite sure that you will find **a** suitable accommodation ( )
6.	**Countability of Noun** When a noun can take only one form because it is uncountable, but an invalid pluralized form has been used, it is a Countability of Noun error.	[However] we have a nice garden, where in summer our students often spend their break or do their **homeworks**. (homework) It is known that **radiations** coming from a mobile phone can be [heart damaging]. (radiation)
7.	**Countability of Quantifier** When a quantifier is incorrect because of the countability of the noun to which it refers, it is a Countability of Quantifier error.	... the groups should have included **less** people. (fewer) Not only had it **small** leg room but also the safety belts were out of order. (little)
8.	**Derivation of Adjective** Where an adjective resembles, or includes the stem of, a valid adjective but has been incorrectly derived, usually because it has been given an incorrect affix, it is a Derivation of Adjective error.	To sum up, the tour was a **completely** disaster. (complete)

#	Category and description	Examples
9.	**Derivation of Adverb** Where an adverb resembles, or includes the stem of, a valid adverb but has been incorrectly derived, usually because it has been given an incorrect affix, it is a Derivation of Adverb error.	*...in one word, everything is so simple. (in a word)* *As recent as ten years ago, ... (recently)*
10.	**Derivation of Anaphor** Where an anaphor resembles, or includes the stem of, a valid anaphor but has been incorrectly derived, usually because it has been given an incorrect affix, it is a Derivation of Anaphor error.	*It's a pleasure to work with your for another year. (you)* *All together I would like to stress that the conference was very useful to my and helped me in my career. (me)*
11.	**Derivation of Conjunction (Link Word)** Where a conjunction / link word resembles, or includes the stem of, a valid word but has been incorrectly derived, usually because it has been given an incorrect affix, it is a Derivation of Conjunction error.	*By the time my grandparents were young, there were so many norms and rules and restrictions about how you should look: ... (When)* *...people seem not to have enough time to have some rest, not even to cook! (or)*
12.	**Derivation of Determiner** Where a determiner resembles, or includes the stem of, a valid determiner but has been incorrectly derived, usually because it has been given an incorrect affix, it is a Derivation of Determiner error.	*You report intentionally hurt our feelings ... (Your)* *It name is "Superstar". (Its)*
13.	**Derivation of Preposition** Where a preposition resembles, or includes the stem of, a valid preposition but has been incorrectly derived, usually because it has been given an incorrect affix, it is a Derivation of Preposition error.	*Not like a lot of other banks which hurried [into] lau[n]ching their own product... (Unlike)* *Related to the gym, it [cannot] be closed. (Regarding)*
14.	**Derivation of Quantifier** Where a quantifier resembles, or includes the stem of, a valid quantifier but has been incorrectly derived, usually because it has been given an incorrect affix, it is a Derivation of Quantifier error.	*Although I like computers the company which I used to work for didn't give a special [training]... (any)* *... apparently these days there was none show prepared. (no)*
15.	**Form of Adjective** When a valid form of the adjective (positive, comparative, or superlative) has been used but is the wrong form in the context.	*They said that it was the worse strike ever and that it would last a long time. (worst)* *This is the most simple way for our company to reach the top. (simplest)*
16.	**Form of Anaphor** When a valid form of the pronoun (uninflected or inflected form) has been used but it is the wrong form in the context.	*Everyone wants to achieve higher marks than other. (others)* *...a [cheque] from yourself... (you)*
17.	**Form of Determiner** When the articles 'a' and 'an' are confused.	*These three things are the ones that we consider to be of more importance.(most)* *We are fortunate enough to have a government which regards education as a important thing ... (an)*

18.	**Form of Noun** When the singular form of a noun is used where the context demands a plural, or vice versa.	Despite having to pay such hefty **amount** to drive, ... (amounts) Thank you for your **considerations.** (consideration)
19.	**Form of Quantifier** When a valid form of the quantifier (singular or plural form) has been used but is the wrong form in the context.	...movies and copies from **another** countries...(other) Just try to get out more and use modern technology as **less** as possible. (little)
20.	**Inflection of Adjective** When the learner has created a feasible but non-valid inflected form of the adjective. This code also covers instances where learners mistakenly make adjectives agree with the nouns they modify.	...which resulted in **huges** queues. (huge) ...you'll see that their lifestyle was much quieter, **simpler** and slower. (simpler)
21.	**Inflection of Adverb** When the learner has created a feasible but non-valid inflected form of the adverb.	Nobody from my group got **farer** from the Hotel's [lobby] ... (further) ...we need to do things **quicklier** than before ...(more quickly)
22.	**Inflection of Anaphor** When the learner has created a feasible but non-valid inflected form of the anaphor.	These services are only the first package of **severals**[ ] which will follow over the next 12 months[.] (several) Each time **everythings** changes. (everything)
23.	**Inflection of Determiner** When the learner has created a feasible but non-valid inflected form of the determiner, usually because of a mistaken belief that the determiner must agree in number with the noun which it precedes.	...because **ours** classes finish at seven o'clock p.m. (our) ...and then travel to **your's** friend[s'] countries. (your)
24.	**Missing Adverb** When a sentence or construction requires an adverb for completeness and that word has been omitted.	These cases **are few** and far between to take place in real life. (are too few) Furthermore, because of this incident I missed a chance to see my friend, when we **had arranged** our meeting. (had already arranged)
25.	**Missing Determiner** When a sentence or construction requires a determiner for completeness and that word has been omitted.	... it was really very difficult to find some kind of entertainment except watching TV **in rooms.** (in our rooms) The quality **of lunch** you provided was the next problem. (of the lunch)
26.	**Missing Noun** When a sentence or construction requires a noun for completeness and that word has been omitted.	By the way, the **train will** take you about three and a half hours and cost around 40 marks. (train journey will) ... the department has a very **limited office** equipment. (limited amount of office)
27.	**Missing Verb**	

#	Type / Description	Example
	When a sentence or construction requires a verb for completeness and that word has been omitted.	...the organisers had never promised that 45 **stalls be** present ... (stalls would be) This would not only reduce the traffic con[g]estion problem in the city, **it also enable** people to own a car ... (it will also)
28.	**Replace Adverb** When a valid adverb in the language has been used and it is the correct part of speech but not the correct adverb.	...the one about tourism was **high** above standard. (well) I have noticed that breakfast is no **more** a family thing. (longer)
29.	**Replace Quantifier** When a valid quantifier in the language has been used and it is the correct part of speech but not the correct word.	I hope my suggestions will be of **any** use to you. (some) Moreover, there are only two telephones for **all of the** department. (the whole of the)
30.	**Replace Verb** When a valid verb in the language has been used and it is the correct part of speech but not the correct verb.	In this context, we cannot **oversee** the changing role of women ... (overlook) It **depends** from one person to the next. (differs)
31.	**Unnecessary Conjunction (Link Word)** When an unnecessary extra conjunction / link word has been used in a sentence or construction in such a way that it makes the sentence or construction incorrect.	We have got enough time to do something for entertainment, such as doing sports, watching movies, **and** etc. **As** I know two people [in] the group would like to go for long walks.
32.	**Unnecessary Determiner** When an unnecessary extra determiner has been used in a sentence or construction in such a way that it makes the sentence or construction incorrect.	The host families are very kind and hospitable and they will also give you **a very good** food.
33.	**Unnecessary Noun** When an unnecessary extra noun has been used in a sentence or construction in such a way that it makes the sentence or construction incorrect.	So, we suggest a visit to a nightclub **disco** where we can hear some modern music and dance. Both hotels can provide full board **meals** and nice comfortable rooms.
34.	**Unnecessary Quantifier** When an unnecessary extra quantifier has been used in a sentence or construction in such a way that it makes the sentence or construction incorrect.	After **all** this terrible experience with your company ... We are hardly [ever] at a loss for words when meeting **some** clients.
35.	**Complex Error** Where an ungrammatical and incomprehensible string of words are written and it is impossible to correct them, it is a Complex Error.	Pacino's career experience seems to be **exploding at his maximum**. And finally I would like to **make again note** we couldn't take an exam. (mention again that)

d) An example of how a grammatical feature develops in learner language across the CEFR levels

Table 12: The development of usage of uncountable nouns across CEFR levels

	Typical use of structure/feature	Examples	Most common errors and CLC examples at the level
A1	Can use some uncountable nouns correctly: *information, money, time, work* with a limited range of determiners: *some, a lot of, the*	*She usually does the housework on Sunday.* *I have a lot of work in college.* *I have got a lot of information about the shops near our area.*	homeworks, breads, houseworks • *I don't like homeworks.* • *She looks after the children every day. Then she does the houseworks in her house.* • *You can buy all types of fruits, sauces, breads.*
A2	Can use an increasing range of common uncountable nouns correctly with an increasing small range of items before a noun: *a lot of, some, more, the important*	*Can you help me? I need more information about this course.*	informations, musics, papers, loves, homeworks • *I have some informations about the art class.* • *You have to bring a pen and some papers.*
B1	Can use an increasing repertoire of uncountable nouns correctly with a greater range of items before the noun: *the, further, more, some, this, interesting, detailed, useful* Errors persist with the most common items.	*Our present accommodation is too small for us.* *If you need further information just ask.*	informations, equipments, homeworks, advices, furnitures, countrysides, works, softwares, trainings • *I'm looking forward to hearing from you if you require further informations.* • *250 of them are very young and they need a lot of trainings.*
B2	Can use an increasing repertoire of uncountable nouns correctly with a greater range of items before the noun: *the, further, more, some, this, interesting, detailed, useful, following, no* Errors increase with the most common items.	*A better way of spending the money is to build a new fitness club.* *I would like to know if I need extra money.* *Tourism has an inevitable fact of bringing a huge sum of money to local businesses.*	informations, advices, equipments, transports, knowledges, works, spending, trainings, homeworks, researches, furnitures, behaviours, damages • *Please send me all the informations and the application.* • *I think both of them have reasonable prices for their foods.*
C1	Can use an increasing repertoire of uncountable nouns correctly with a greater range of items before the noun: *the, further, more, some, this, interesting, detailed, useful, additional, essential, up-to-date*	*For more detailed information you may contact me on...* *They have already been equipped with the necessary hardware and software and have access to the internet and our database.*	informations, transports, trainings, equipments, advices, knowledges, softwares, researches, furnitures, tuitions, spendings, accommodations, feedbacks, congestions • *For readers that are less keen on sports, an article containing some simple advices on how to keep fit will*

	Errors increase with the most common items and with determiners.	*The only criticism we can make is the amount of* ***equipment****.*	• *be more pleasant to read.* • *I had a small problem with* ***transports****, because the nearest bus stop was about two miles far away.* • *The three computers were acquired in order to replace the out-of-date* ***equipments*** *that were used by the General Manager.*
C2	Can use an increasing repertoire of uncountable nouns correctly with a greater range of items before the noun: *the, further, more, some, this, interesting, detailed, useful, additional, essential, up-to-date, enough, provide* Errors decrease but persist with the most common items.	*Research and development has prevented many of us from illnesses…* *It is for these people that a magazine such as yours should give extensive* ***advice****.* *In addition, there has also been significant* ***damage*** *to our houses: walls and ceilings cracking, roof tiles coming loose, and occasionally window panes being shattered!*	*advices, behaviours, informations, researches* • *Reading these books are very interesting, as* ***many advices are*** *hidden inside.* • *Even though some of our childhood influences may lead us to negative* ***behaviours*** *as adults, such as being messy, it might be a good idea to be thankful for the good influences.* • *You can even exchange your teaching and learning experiences or* ***informations*** *with other schools through computers.* • *All* ***informations*** *in this statement* ***are*** *provided with my best knowledge.* • *Many* ***researches*** *have been made already in order to find out if other planets, like the moon, Mars or Venus, are suitable for human life.*

These findings on grammar are still tentative and suggest future research directions, including extending the concept of "criteriality" to the analysis to other samples of learner language and in particular to speech. Spoken language data is being collected and corpora are being built with the necessary computational tools to enable the research to be extended in that direction. In addition, a focus on the C-levels continues with the collection of academic English writing samples. Work is underway to produce an *English Grammar Profile* resource that both documents and presents these findings in a user-friendly way, in line with the general English Profile approach.

3 The English Functions Profile

This research strand re-examines the use of language functions in defining language constructs across the ability range and in setting learning objectives which are sensitive to the proficiency level of learners. This part of the EP Programme is primarily concerned with the *input* to the learners: in other words, the functions that provide the learning objectives as part of a communicative syllabus at each level of the CEFR, and the reading texts which are judged to be suitable for different levels and which are presented to the learners for pedagogic purposes. This research work is based on the *functional-notional approach* which underpins the CEFR and which led to the use of Can Do statements within its Global Scale and bank of Illustrative Scales. English Profile suggests practical ways in which the existing Can Do statements in the CEFR can be expanded and refined by providing additional detail with reference to contexts of use which are particularly relevant to learners of English.

Language functions aim to capture not what learners know about a language, or what aspects of the language they are able to use (i.e. the lexico-grammatical aspects of language), but rather they intend to convey how learners use the language: what they *can do* with it in social contexts. When investigating functions, in addition to the CEFR, English Profile builds on the Threshold Series or *T-series* (van Ek 1975; van Ek & Trim 1998a, 1998b, 2001; Trim 2009). The T-series was instrumental in promoting a communicative approach to language teaching from the 1970s onwards and covered proficiency levels that are now associated with CEFR A2 (Waystage), B1 (Threshold), and B2 (Vantage). An A1 (Breakthrough) specification by John Trim has also been made available electronically for the EP Programme and all T-series publications are now freely accessible from the English Profile website.

As Green (2011) points out, the CEFR's approach was clearly influenced by Hymes (1972) and a conceptualization of *communicative competence* which comprises both a "linguistic" and a "sociolinguistic" element. The sociolinguistic dimension includes the dynamic interaction that occurs between the context and the discourse produced. The interaction between context and the cognitive processes of the learner underpins a *socio-cognitive approach* to learning, teaching and assessment on which the CEFR itself is based, and which has been important in developing the English Profile Programme of research.

English Profile proposes a new, generative, Can Do model consisting of five elements which can be linked to specific contexts and which can lead to a more detailed and technical definition of the CEFR levels for English (see definition alongside).

These Can Do statements are intended to help and encourage practitioners to develop and validate Can Do statements to fit more appropriately with their own context and purposes, while at the same time, finding a way to relate them coherently to the common framework. This aligns with the CEFR's functions which are 'widely interpretable … based on the everyday expressions used by teachers' (Green 2011: 14).

Defining *generative Can Do statements:*

'frames setting out how the elements of the CEFR model may interact in shaping the difficulty of defined language activities and tasks.' (Green 2011: 160)

'Can Do statements … should provide enough information to guide users in carrying out a variety of purposes. They should offer sufficient detail to inform materials writers and test developers (who need to operationalise the general CEFR framework through specific, contextualized tasks); but they should also support briefer summary statements that might communicate suitably general information to others, such as the users of test results … The statements should also serve as a link between these different purposes: users should be able to trace the ways in which the elements of the framework are interpreted and represented in the specific demands made of learners in using language to carry out a task.'
(Green 2011: 47)

The component elements proposed for the new Can Do statements include the following:

Activity: Can…	The social act (function) or related sequence of acts (activity) that the learner might be expected to accomplish by means of the language, i.e. what the learner can do
Theme/Topic/Setting: Concerned with…	The themes, topics and settings in relation to which the learner might be expected to perform. In the CEFR, applicable themes are grouped under the four domains: educational, public, professional and personal, i.e. what the interaction is concerned with
Input text: Based on…	The nature of the text that the learner might be required to process as a basis for his or her own contribution or to demonstrate his or her comprehension, i.e. what the communication is based on
Output text: Producing…	The nature of the text that the learner might be expected to produce or participate in producing to demonstrate (a specified degree of) understanding or to accomplish a task
Qualities: How well?	The qualities that the learner would be expected to demonstrate in carrying out language activities. For production, these qualities are grouped under the CEFR headings of Linguistic, Pragmatic, Sociolinguistic and Strategic competences and would answer the question How well?
Restrictions: Provided that…	Physical or social conditions and constraints under which the learner would be expected to perform, i.e. Provided that …

In terms of language functions, English Profile has focused on getting a better understanding of English used at the C levels because these are the least well defined in the CEFR and were not covered by the T-series. In keeping with the evidence-led stance of the English Profile Programme, these Can Do statements have a strong empirical basis. English Profile researchers collected data in the form of language learning materials which are aligned to the CEFR and are in use around the world. The range of sources that have informed the English Profile Can Do statements to date include:

- CEFR illustrative scales
- bestselling international textbooks and related support materials from different publishers including Cambridge University Press
- examination handbooks from Cambridge ESOL
- curriculum and syllabus documents sourced by English Profile partners such as the British Council, English UK and others
- the *Bank of descriptors for self-assessment in European Language Portfolios* (Lenz & Schneider 2004)
- online publications by educational institutions such as test specifications and handbooks, proficiency scales and textbook support materials
- additional non publicly available materials from various educational contexts.

Through the EP Network (see Sections 7 and 9), the research team consulted widely on the development of the new Can Dos for the C levels. The inventory of these refined and contextualised functions is given in Green 2011 in the form of sets of English Profile Can Do statements. These Can Do Statements are reproduced here and are divided by mode (spoken or written) and whether relevant for interaction, production or reception.

English Profile Can Do statements for the C levels

Table 13: Can Do statements for Spoken Interaction at the C levels

Activity Can...	Topic/Setting Concerned with...	Input Based on...	Text Producing...	Qualities – how well? Linguistic/ Pragmatic/ Sociolinguistic/ (X Strategic)	Restrictions Provided that...
1. account for and sustain his/her opinions	complex technical/ abstract topics		[spoken interaction] discussion	P9. providing relevant explanations, arguments and comments L2. is readily comprehensible L3. intonation is used to support meaning L4. uses a range of grammatical forms and vocabulary appropriate to the audience, topic and social context S2. adjusts level of formality and style of speech to suit social context: formal, informal, colloquial P10. contributions are relevant and are integrated into the flow of the discourse using linking words, repetition of words used by other participants	
2. discuss the nature and relative merits of particular choices	goods or services, procedures, courses of action		[spoken interaction]	L2. is readily comprehensible L3. intonation is used to support meaning L4. uses a range of grammatical forms and vocabulary appropriate to the audience, topic and social context S2. adjusts level of formality and style of speech to suit social context: formal, informal, colloquial P10. contributions are relevant and are integrated into the flow of the discourse using linking words, repetition of words used by other participants X2. repairs interaction as necessary	
3. follow and contribute to complex interactions between third parties	abstract, complex unfamiliar topics		[spoken interaction] group discussion	P11. turn taking is natural P10. contributions are relevant and are integrated into the flow of the discourse using linking words, repetition of words used by other participants P8. conveys ideas with some precision, does not resort to simplification L2. is readily comprehensible L3. intonation is used to support meaning L4. uses a range of grammatical forms and vocabulary appropriate to the audience, topic and social context	

No.	Can-do statement	Topic domain	Text type	Performance criteria
				S2. adjusts level of formality and style of speech to suit social context: formal, informal, colloquial X2. repairs interaction as necessary
4.	evaluate, restate and challenge	matters within their academic or professional competence	contributions from other participant	[spoken interaction] P10. contributions are relevant and are integrated into the flow of the discourse using linking words, repetition of words used by other participants. L2. is readily comprehensible L3. intonation is used to support meaning L4. uses a range of grammatical forms and vocabulary appropriate to the audience, topic and social context S2. adjusts level of formality and style of speech to suit social context: formal, informal, colloquial
5.	frame critical remarks or express strong disagreement in such a way as to minimise any offence	familiar topics/ matters within their academic or professional competence		[spoken interaction] P12. remarks are relevant P10. contributions are relevant and are integrated into the flow of the discourse using linking words, repetition of words used by other participants. S2. adjusts level of formality and style of speech to suit social context: formal, informal, colloquial L2. is readily comprehensible; L3. intonation is used to support meaning L4. uses a range of grammatical forms and vocabulary appropriate to the audience, topic and social context X2. repairs interaction as necessary
6.	express sympathy or condolence, enquire into the causes of unhappiness or sadness and offer comfort	personal relationships/ counselling roles		[spoken interaction] conversation L2. is readily comprehensible L6. uses intonation and word choice to express mood, distinguishing between shades of feeling L4. uses a range of grammatical forms and vocabulary appropriate to the audience, topic and social context S2. adjusts level of formality and style of speech to suit social context: formal, informal, colloquial
7.	discuss work with colleagues	matters within their academic or professional competence		[spoken interaction] L5. using appropriate technical terminology L2. is readily comprehensible L3. intonation is used to support meaning P8. conveys ideas with some precision, does not resort to simplification L4. uses a range of grammatical forms and vocabulary appropriate to the audience, topic and social context

				S2. adjusts level of formality and style of speech to suit social context: formal, informal, colloquial P10. contributions are relevant and are integrated into the flow of the discourse using linking words, repetition of words used by other participants X2. repairs interaction as necessary
8.	indicate levels of willingness or reluctance and state conditions when agreeing to requests or granting permission	requests	[spoken interaction]	L2. is readily comprehensible L3. intonation is used to support meaning L4. uses a range of grammatical forms and vocabulary appropriate to the audience, topic and social context S2. adjusts level of formality and style of speech to suit social context: formal, informal, colloquial
9.	invite participation, introduces issues, manages contributions	meetings on matters within their academic or professional competence	[spoken interaction]	P12. plays a leading part in discussion or negotiation L2. is readily comprehensible L3. intonation is used to support meaning L4. uses a range of grammatical forms and vocabulary appropriate to the audience, topic and social context S2. adjusts level of formality and style of speech to suit social context: formal, informal, colloquial X2. repairs interaction as necessary
10.	summarise and evaluate main points in discussion	meetings on matters within their academic or professional competence	[spoken interaction] extended discussion or negotiation	P13. able to gain and hold floor P14. summary is accurate L2. is readily comprehensible L3. intonation is used to support meaning L4. uses a range of grammatical forms and vocabulary appropriate to the audience, topic and social context S2. adjusts level of formality and style of speech to suit social context: formal, informal, colloquial
11.	enquire about and negotiate special treatment	prices and conditions of sale - rates, terms and conditions	[spoken interaction] extended negotiation	L5. using appropriate technical terminology L2. is readily comprehensible L3. intonation is used to support meaning P8. conveys ideas with some precision, does not resort to simplification L4. uses a range of grammatical forms and vocabulary appropriate to the audience, topic and social context S2. adjusts level of formality and style of speech to suit social context: formal, informal, colloquial X2. repairs interaction as necessary

#	Function	Topic	Context	Criteria	Notes
12.	complain and negotiate redress	poor service e.g. returning faulty, inappropriate or unwanted goods and negotiating for a replacement or refund	[spoken interaction] extended negotiation	L5. using appropriate technical terminology L2. is readily comprehensible L3. intonation is used to support meaning P8. conveys ideas with some precision, does not resort to simplification L4. uses a range of grammatical forms and vocabulary appropriate to the audience, topic and social context S2. adjusts level of formality and style of speech to suit social context: formal, informal, colloquial X2. repairs interaction as necessary	
13.	apologise and ask for understanding of their position	legal, regulatory matters: in case of infringements of regulations	[spoken interaction] exchanges with officials	L2. is readily comprehensible L3. intonation is used to support meaning L4. uses a range of grammatical forms and vocabulary appropriate to the audience, topic and social context S2. adjusts level of formality and style of speech to suit social context: formal, informal, colloquial X2. repairs interaction as necessary	infringements are minor
14.	express regrets and negative wishes or intentions		[spoken interaction]	L2. is readily comprehensible L3. intonation is used to support meaning L4. uses a range of grammatical forms and vocabulary appropriate to the audience, topic and social context S2. adjusts level of formality and style of speech to suit social context: formal, informal, colloquial X2. repairs interaction as necessary	
15.	establish solidarity with interlocutors through sympathetic questioning and expressions of agreement	complaints about third parties/ conditions	[spoken interaction] informal conversation (with friends)	P15. contributions are integrated into the flow of the discourse using linking words, repetition of words used by other participants S2. adjusts level of formality and style of speech to suit social context: formal, informal, colloquial L2. is readily comprehensible L6. uses intonation and word choice to express mood, distinguishing between shades of feeling L4. uses a range of grammatical forms and vocabulary appropriate to the audience, topic and social context X2. repairs interaction as necessary	

Activity Can...	Topic/Setting Concerned with...	Input text Based on...	Output text Producing...	Qualities – how well? Linguistic/ Pragmatic/ Sociolinguistic/ (X Strategic)	Restrictions Provided that...
16. ask for explanation or clarification and negotiate understanding	complex, abstract ideas academic/ professional matters		[spoken interaction] discussion	L2. is readily comprehensible L3. intonation is used to support meaning L4. uses a range of grammatical forms and vocabulary appropriate to the audience, topic and social context S2. adjusts level of formality and style of speech to suit social context: formal, informal, colloquial X2. repairs interaction as necessary	
17. negotiate a course of action with a partner or group, reporting on what others have said, summarizing, elaborating and weighing up multiple points of view			[spoken interaction] discussion	P15. contributions are integrated into the flow of the discourse using linking words, repetition of words used by other participants L2. is readily comprehensible L3. intonation is used to support meaning L4. uses a range of grammatical forms and vocabulary appropriate to the audience, topic and social context S2. adjusts level of formality and style of speech to suit social context: formal, informal, colloquial X2. repairs interaction as necessary	

Table 14: Can Do statements for Spoken Production at the C levels

Activity Can...	Topic/Setting Concerned with...	Input text Based on...	Output text Producing...	Qualities – how well? Linguistic/ Pragmatic/ Sociolinguistic/ (X Strategic)	Restrictions Provided that...
18. interpret specialist topics to the layperson	complex technical topics		[spoken production] addressing audiences	L1. able to speak at length as required L2. is readily comprehensible L3. intonation is used to support meaning L4. good command of non-technical circumlocution, idiomatic expressions and colloquialisms L4. uses a range of grammatical forms and vocabulary appropriate to the audience, topic and social context P2. manipulates the order of elements to control information focus S1. makes topic accessible to the layperson S2. adjusts level of formality and style of speech to suit social context: formal, informal, colloquial S3 consistent register X1. checks comprehension	topics relate to his/her field of interest presentation prepared independently
19. qualify assertions	complex technical/ abstract topics		[spoken production] addressing audiences	L1. able to speak at length as required L2. is readily comprehensible L3. intonation is used to support meaning	

			Descriptors	
			L4. uses a range of grammatical forms and vocabulary appropriate to the audience, topic and social context L5. indicating levels of confidence or uncertainty P1. clear P3. demonstrates flexibility and control of nuances P7. well-structured and developed	topics relate to his/her field of interest
20. define or specify	complex technical/ abstract topics	[spoken production] addressing audiences	L2. is readily comprehensible L3. intonation is used to support meaning L4. uses a range of grammatical forms and vocabulary appropriate to the audience, topic and social context P1. clear P2. demonstrates flexibility and control of nuances P4. in detail, distinguishing between objects or concepts that closely resemble each other P5. at length P7. well-structured and developed	
21. give instruction	series of complex professional/ academic procedures	[spoken production] [sustained monologue]	L2. is readily comprehensible L3. intonation is used to support meaning L4. uses a range of grammatical forms and vocabulary appropriate to the interlocutor(s), topic and social context L5. using appropriate technical terminology P1. clear P5. at length P6. detailed P7. well-structured and developed P8. conveys ideas with some precision, does not resort to simplification X1. checks comprehension as necessary	instructions to a work colleague or student with some technical knowledge
22. speculate or hypothesise, comparing and evaluating a number of possible developments		[spoken production] [sustained monologue]	P1. clear P6. detailed P5. a t length P7. well-structured and developed L2. is readily comprehensible L3. intonation is used to support meaning P8. conveys ideas with some precision, does not resort to simplification L4. uses a range of grammatical forms and vocabulary appropriate to the interlocutor(s), topic and social context	

Table 15: English Profile Can Do statements for Spoken Reception at the C levels

Activity Can…	Text When listening to…	Text characteristics Characterised by…	Topic/Setting Concerned with…	Qualities How well?	Restrictions Provided that…
23. make notes to extract and reconstruct the main points and key supporting details	presentations, lectures or documentary broadcasts	extended monologue	complex public, academic or professional topics involving detailed propositional information that is new to the listener and includes abstract concepts	produces accurate and detailed summary or text level representation of factual content summarises the main points accurately responds to questions of detail	standard accents, familiar to the listener rate of speech is natural
24. extract the gist and distinguish between opinions	informal meetings and discussions	multi-participant discussion with non-linear organisation, frequent colloquialisms and overlapping turns	personal or public topics of general interest	identifies the main points being made by participants identifies all areas of agreement and disagreement between participants	conversation is animated – at a fast natural rate voices are easily differentiated or audio is supported by images language is standard, but a range of accents are used
25. identify the emotions or attitudes of speakers	informal meetings and discussions	multi-participant discussion marked by non-linear organisation, colloquialisms and overlapping turns	personal, public, academic, professional topics	accurately identifies the attitudes or emotions conveyed implicitly by stress, pitch and intonation, lexical choices	conversation is animated – at a fast natural rate voices are easily differentiated or audio is supported by images language is standard, but a range of accents are used
26. extract gist, detail, purposes and main points	formal discussions on academic, public or professional topics	dialogues, multi-participant discussion with formal turn taking and organisation – may be mediated by a chairperson	complex public, academic or professional topics involving detailed propositional information that is new to the listener: facts, definitions	produces accurate and detailed summary or text level representation of factual content, showing relationships between ideas accurately responds to questions of detail	contributions are clearly presented – intended for an audience as well as fellow participants standard accent, familiar to the listener
27. extract, select and	multiple sources:	variety of spoken	complex, abstract	extracts and evaluates information and opinions	standard accent,

integrate detailed information required to carry out related tasks		text types: extended monologues, multi-participant discussions	personal, public, academic, professional topics	from different sources integrates these in preparation for a report, essay or position paper etc.	familiar to the listener discussions are formal and structured
28. identify, analyse and evaluate the use of interactive spoken language for persuasion	formal debates, interviews, business interactions, situations of personal or public conflict	dialogues, multi-participant discussions	personal, public, academic, professional topics	identifies how linguistic resources (stress, pitch and intonation, lexical choices) are used by participants to resolve conflict, build consensus, promote views etc. and evaluates the success of these strategies in an interaction	in dialogues/ discussions, voices are easily differentiated or audio is supported by images language is standard, but a range of accents are used
29. evaluate presentations in relation to their purpose and audience	presentations, speeches,	extended monologue	personal, public, academic, professional topics	identifies speaker purpose and intended audience evaluates the use of language in relation to these: suggests improvements	language is standard, but a range of accents may be used rate of speech is natural
30. integrate information and detailed instructions to carry out complex tasks involving multiple elements	multiple sources	extended monologues, dialogues	personal, public, academic, professional topics	brings together information from different sources to describe a task to be carried out and steps to complete the task	language is standard, but a range of accents may be used rate of speech is natural unfamiliar process or procedure the context is familiar – personal to the listener or within the listener's academic/ professional field

Table 16: English Profile Can Do statements for Written Interaction at the C levels

Activity Can...	Topic/Setting Concerned with...	Input text Based on...	Output text Producing...	Qualities – how well? Linguistic/ Pragmatic/ Sociolinguistic	Restrictions Provided that...
31. Write in support of a candidate for a job or award	personal or professional	resume	[written interaction] letter/ email of reference	L1. maintains high levels of linguistic accuracy over extended text L2. makes effective use of linguistic modality to signal the strength of claim, argument, or position L5. complex P1. clear P2. appropriate and effective logical structure which helps the reader to find significant points P6. reaches a position or conclusion P7. well-structured and developed P9. logical P10. uses conventional elements of genre structure S1. style appropriate to the genre adopted and to the reader in mind	
32. Write a persuasive application	academic or professional employment	personal information	[written production] job/ study application letter/ email	L1. maintains high levels of linguistic accuracy over extended text L2. makes effective use of linguistic modality to signal the strength of claim, argument, or position L5. complex P1. clear P2. appropriate and effective logical structure which helps the reader to find significant points P6. reaches a position or conclusion P7. well-structured and developed P9. logical P10. uses conventional elements of genre structure S1. style appropriate to the genre adopted and to the reader in mind	employment/ study within his/her field of interest
33. Evaluate, restate and challenge an argument	professional or academic	online discussion	[written interaction] online discussion	L1. maintains high levels of linguistic accuracy over extended sequence of turns P1. clear P10. uses conventional elements of genre structure S1. style appropriate to the genre adopted and to the reader in mind X2. repairs interaction as necessary	in real time
34. Ask for	complex, abstract		[written interaction]	L1. maintains high levels of linguistic accuracy over	in real time

explanation or clarification and negotiate understanding	ideas academic/ professional matters	online discussion	extended sequence of turns P1. clear P10. uses conventional elements of genre structure S1. style appropriate to the genre adopted and to the reader in mind X2. repairs interaction as necessary
35. express sympathy or condolence and offer comfort	sensitive personal matters	news of bereavement/ divorce	[written interaction] letter of sympathy, condolence	L1. maintains high levels of linguistic accuracy over extended text L2. makes effective use of linguistic modality to signal the strength of claim, argument, or position L5. complex P1. clear P7. well-structured and developed P9. logical P10. uses conventional elements of genre structure S1. style appropriate to the genre adopted and to the reader in mind

Table 17: English Profile Can Do statements for Written Production at the C levels

Activity Can...	Topic/Setting Concerned with...	Input text Based on...	Output text Producing...	Qualities – how well? Linguistic/ Pragmatic/ Sociolinguistic	Restrictions Provided that...
36. incorporate information drawn from the work of others into his/her own text	complex academic or professional topics	professional or academic texts	[written production] reports, articles or essays	P1. clear / P2. appropriate and effective logical structure which helps the reader to find significant points / P3. without infringing conventionally-accepted academic/professional standards of the use of others' work	
37. set out multiple perspectives on an intellectual issue	complex academic or professional topics	professional or academic texts	[written production] reports, articles or essays	L1. maintains high levels of linguistic accuracy over extended text / L2. makes effective use of linguistic modality to signal the strength of claim, argument, or position / P4. makes clear author's own stance on the issue / P5. clearly distinguishes own ideas and opinions from those of (multiple) sources / S1. style appropriate to the genre adopted	topic within his/her field of interest
38. describe and interpret	complex academic or professional topics	empirical data from research	[written production] reports, articles or essays	L1. maintains high levels of linguistic accuracy over extended text / L3. at length (500 words plus) / P1. clear / P6. reaches a position or conclusion / P7. well-structured and developed / S1. style appropriate to the genre adopted	
39. present specialist material	complex academic or professional topics	research or professional/ academic texts	[written production] reports, articles or essays for a general audience	L1. maintains high levels of linguistic accuracy over extended text / L3. at length (500 words plus) / L4. uses suitably non-technical words and phrases / P1. clear / P7. well-structured and developed / S1. style appropriate to the genre adopted and to the reader in mind / S2. accessible to an audience that is not familiar with the topic	topic within his/her field of interest opportunities for redrafting and revision
40. define or specify	complex academic or professional topics	Personal experiences/ texts	[written production] reports, articles or essays	L3. at length (500 words plus) / P8. in detail, distinguishing between objects or concepts that closely resemble each other / L1. maintains high levels of linguistic accuracy over	

			extended text P1. clear L5. complex P9. logical P7. well-structured and developed	topic within his/her field of interest opportunities for redrafting and revision
41. write an introduction	complex academic or professional topics	Co-text [written production] longer reports, articles or dissertations	L1. maintains high levels of linguistic accuracy over extended text L3. at length (500 words plus) L5. complex P1. clear P7. well-structured and developed P9. logical P10. uses conventional elements of genre structure P11. presenting key elements to be developed in succeeding text	topic within his/her field of interest opportunities for redrafting and revision
42. write a conclusion	complex academic or professional topics	Co-text [written production] longer reports, articles or dissertations	L1. maintains high levels of linguistic accuracy over extended text L3. at length (500 words plus) L5. complex P1. clear P7. well-structured and developed P9. logical P10. uses conventional elements of genre structure P12. reviews and summarises an extended exposition or argument	

Table 18: English Profile Can Do statements for Written Reception at the C levels

Activity Can…	Text When reading…	Text characteristics Characterised by…	Topic/Setting Concerned with…	Qualities – how well?	Restrictions Provided that…
43. demonstrate comprehensive understanding	articles in serious newspapers or magazines; reference books; specialised academic/ professional publications	lengthy, complex sentences infrequent, sometimes technical vocabulary formal register	complex public, academic or professional topics conveying detailed propositional information that is new to the reader: facts, definitions	produces accurate and detailed summary or text level representation of factual content, showing relationships between ideas accurately responds to questions of detail	access to reference tools opportunities for re-reading texts addressed to general educated readership OR texts addressed to professional/ academic community to which learner belongs
44. integrate ideas across texts: compare, contrast, synthesize	articles in serious newspapers or magazines; reference books; specialised academic/ professional publications	lengthy, complex sentences infrequent, sometimes technical vocabulary formal register	complex, public, academic or professional topics including abstract ideas	identifies all main areas of agreement and disagreement across texts selects elements from two or more texts to construct a balanced response to a question	access to reference tools opportunities for re-reading texts addressed to general educated readership OR texts addressed to professional/ academic community to which learner belongs
45. demonstrate understanding of implicit attitudes and opinions	articles in serious newspapers or magazines; reference books; specialised academic/ professional publications	linguistic means for conveying attitude include use of metaphor/ marked syntax/ lexical connotation Texts characterised by: lengthy, complex sentences infrequent, sometimes	complex public, academic or professional topics; conveying attitudes and opinions	accurately summarises the views of the writer	access to reference tools opportunities for re-reading texts addressed to general educated readership OR texts addressed to professional/

	Text types	Features	Topics	Performance	Conditions
		technical vocabulary formal register			academic community to which learner belongs
46. demonstrate comprehensive understanding	personal messages in informal letters, emails etc. informal articles, weblogs etc.	informal register colloquial expressions elliptical cohesion cultural references	complex public, academic or professional topics conveying detailed propositional information that is new to the reader: facts, definitions	produces accurate and detailed summary or text level representation of factual content, showing relationships between ideas accurately responds to questions of detail	access to reference tools opportunities for re-reading texts addressed to general educated readership OR texts addressed to professional/ academic community to which learner belongs
47. demonstrate understanding of implicit attitudes and opinions	personal messages in informal letters, emails etc. informal articles, weblogs etc.	linguistic means for conveying attitude include use of metaphor/ marked syntax/ lexical connotation informal register colloquial expressions elliptical cohesion cultural references	personal topics; conveying emotions, attitudes and opinions	accurately summarises the views of the writer	access to reference tools opportunities for re-reading
48. critically interpret	literary writings articles in serious newspapers or magazines; specialised academic/ professional publications personal messages in informal letters, emails etc.	formal register lengthy, complex sentences infrequent, sometimes technical vocabulary OR informal register colloquial expressions elliptical cohesion cultural references	personal, public, educational or professional topics	accurately represents the views of the writer and engages critically with them to reach a conclusion	access to reference tools opportunities for re-reading

49. can demonstrate awareness of the impact on the reader of features of style: lexical and grammatical choices, discourse organisation	literary writings articles in serious newspapers or magazines; specialised academic/ professional publications personal messages in informal letters, emails etc.	formal register lengthy, complex sentences infrequent, sometimes technical vocabulary OR informal register colloquial expressions elliptical cohesion cultural references	personal, public, educational or professional topics	successfully identifies genre, tone, purpose, stylistic features and likely impact	access to reference tools opportunities for re-reading
50. demonstrate broad understanding	articles in serious newspapers or magazines; reference books; specialised academic/ professional publications	lengthy, complex sentences infrequent, sometimes technical vocabulary formal register	complex public, academic or professional topics conveying detailed propositional information that is new to the reader: facts, definitions	produces accurate general overview of content reproduces all main ideas accurately responds to questions of gist	under time pressure texts addressed to general educated readership OR texts addressed to professional/ academic community to which learner belongs

Research is now underway to specify language functions at the lower levels (A1 and A2) and additional descriptors are being developed to specify functional objectives for younger learners of English who are studying the language as part of their school curriculum. English Profile researchers also aim to provide learner examples of functional usage for each CEFR level, illustrating how related functions develop across the CEFR levels.

The main thrust of the EP Programme's research on vocabulary has resulted in the English Vocabulary Profile (Capel 2011), an interactive online resource which describes the vocabulary which learners of English know at each level of the CEFR. It provides a large searchable database of detailed information on the words and phrases that are appropriate for learners at each level of the CEFR and is already being used to inform teaching, publishing and assessment activities, with further uses and upgrades planned.

Compiling the English Vocabulary Profile

The English Vocabulary Profile research has been substantially but not exclusively corpus-informed. We have used both the *Cambridge English Corpus*, a 1.2 billion-word collection of written and spoken English, and the *Cambridge Learner Corpus*, a unique corpus of written learner English, currently standing at 45 million words and which includes student writing at all six CEFR levels, and from over 200 countries. Work is now underway to construct a spoken learner corpus of 2 million words, which will further inform future iterations of the English Vocabulary Profile by providing evidence of vocabulary in L2 speech, in both testing and other contexts.

In combination with this corpus evidence, we have monitored a range of classroom-based sources, including wordlists from leading coursebooks, readers' wordlists and the content of vocabulary skills books. We have also referred to the Vocabulary Lists for Cambridge ESOL's Key English Test (KET) and Preliminary English Test (PET) examinations, which have been in use since 1994 and have been regularly updated to reflect language change and patterns of use. Finally, even though it was published thirty years ago, the *Cambridge English Lexicon* by Roland Hindmarsh (1980) has proved invaluable as a checking source, where the language has not evolved over time; it too was organised at meaning level.

Many people will already be familiar with the working title of the project, which was *English Profile Wordlists*. However, the resource is much more than a list of words for each CEFR level. The current English Vocabulary Profile:

- contains words, phrases, phrasal verbs and idioms
- presents the level of each meaning of a word in CEFR order, to suggest learning priorities
- provides detailed dictionary-style entries with clear definitions, grammatical information and guidewords to meanings
- includes audio and written pronunciations
- contains many real examples, from dictionaries and from actual learners at an appropriate level
- can be searched according to different filters, including parts of speech, grammar, usage, topic and affixes
- contains both British English and American English versions which users can easily switch between.

What does it mean to "know" a word?

The core objective of the English Vocabulary Profile project has been to establish which words and phrases are commonly known by learners around the world. What is meant by "know" in this context? We have not attempted to separate receptive competence from productive as, in reality, so much will depend on learning styles and priorities. In general, communicative classrooms in the 21st century provide more consistent opportunities for using new language than a generation ago. Added to that is the unlimited access that most students have to the Internet, where they will be browsing but also actively participating through English. For us, 'knowing' a word is a cumulative process, which implies lifelong learning, as further meanings and uses are acquired.

Take for example the word *know.* The English Vocabulary Profile entry for this word stretches from A1 to C2 level, with figurative and idiomatic uses coming in at the higher CEFR levels, suggesting that there are additional

meanings and phrases containing *know* that are not acquired until the C levels, see the core results for *know* in British English below:

Core results:

- know verb HAVE INFORMATION A1
- know verb ASK FOR INFORMATION A1
- know verb BE ABLE A2
- know verb BE CERTAIN A2
- let sb know A2
- you know B1
- know verb BE FAMILIAR WITH B1
- know verb UNDERSTAND A SUBJECT B1
- get to know sb/sth B1
- I know B1
- as you know B1
- as far as I know B2
- know better (than to do sth) B2
- I know B2
- you never know B2
- before you know it C1
- know sth inside out C1
- know what you are talking about C1
- know verb GUESS CORRECTLY C2
- know best C2
- know of sth/sb B2

Key features of the English Vocabulary Profile

There are several key features of the English Vocabulary Profile, starting with its presentation of different meanings for the same word. The English Vocabulary Profile operates at the level of individual meanings, unlike most vocabulary resources as some meanings are quite distinct from the core meaning of a word, and will be encountered by learners at different CEFR levels. The English Vocabulary Profile team have evaluated each sense of a word, starting from its frequency for first language users and comparing that with learner data through CUP's corpus-informed dictionary database, which is the only monolingual English dictionary resource to flag frequency at sense level.

The EVP also includes extensive information about phrases based on state-of-the-art research on phrasal expressions, for example *make your way*, which has a B2 sense of literally get to a place versus the C2 figurative sense of make your way in a career:

make your way

B2 to get to a place

Dictionary examples:

We slowly made our way down the river.

It's getting late - we should make our way home soon.

> ⊙ **Learner example:**
> *She made her way slowly to the waiting room and patiently waited for the train to arrive.*
>
> First Certificate in English; B2; Azerbaijani

all the way

B2 the complete distance

Dictionary example:

They cycled all the way to London.

> ⊙ **Learner example:**
> *I drove all the way to Bergen in only one day.*
>
> First Certificate in English; B2; Swedish

Similarly, learner knowledge of phrasal verbs is being investigated by an EP Network partner in Japan who is conducting a large-scale test to evaluate the relative difficulty of 100 of the 442 phrasal verbs that are currently included in the A1 to B2 levels. This test, which is being replicated in other regions, will enable us to compare knowledge of phrasal verbs across different first language backgrounds.

Here is an extract from the entry for the word *cool* (showing A1 to B2 level) that illustrates some of meanings available and key phrasal verbs including *cool*:

cool /kuːl/   **Outline View**

▶ **EXCLAMATION** INFORMAL

A2 used when you like something or agree to something
   Dictionary examples:
   You've got your own apartment? Cool!
   "We'll meet at ten, then?" "Cool."
   Learner example:
   That's very exciting. Cool!
                                                            Key English Test; A2; Chinese

▶ **ADJECTIVE**

   GOOD
   A2 INFORMAL good, stylish, or fashionable
   Dictionary examples:
   It was a really cool gig.
   She's really cool.
   Learner example:
   I got this mega cool T-shirt.
                                                            Key English Test; A2; Dutch

   SLIGHTLY COLD
   B1 slightly cold, but not too cold
   Dictionary examples:
   cool water
   a nice cool breeze
   Learner example:
   Sometimes a cool wind blows [in Bodrum].

   > Guide words in capital letters for each meaning help the user navigate through long entries

                                                            Preliminary English Test; B1; Turkish

▶ **VERB** [I or T]

   B2 to become less hot, or to make something become less hot
   Dictionary example:
   Allow the bread to cool before slicing it.
   Learner example:
   In the summer many people were delighted by fountains which cooled the air.
                                                            First Certificate in English; B2; German

cool down/off (sb/sth) or cool (sb/sth) down/off *phrasal verb*

   > Phrasal verbs are listed at the end of an entry

   BECOME LESS HOT
   B2 to become less hot, or to make someone or something become less hot
   Dictionary examples:
   She waited until her coffee had cooled down and then took a sip.
   We went for a swim to cool off.
   Learner example:
   The pool is a very popular place to cool down [in].
                                                            First Certificate in English; B2; Swiss German

   BECOME CALMER
   B2 to become calmer, or to make someone become calmer
   Dictionary example:
   Just leave her to cool off and then talk to her.
   Learner example:
   It seemed to me that you were stressed. If I were you, I would cool down a little.
                                                            First Certificate in English; B2; Swedish

The English Vocabulary Profile also provides information on groups of words with a common root (word families). The different parts of speech in a word family will often be at different CEFR levels according to their frequency.

Another key area of research within the vocabulary strand is the investigation of affixation. Here too, native speaker frequency has played a part in the decision-making process as to CEFR level, along with a consideration of the transparency of the different parts of speech in relation to the base word. When an affix is attached to a word, it may be transparent in meaning – for example, it is not difficult to understand *downloadable* from the verb *download* - but it is less easy to work out the meaning of *changeable*, as in *changeable weather*, formed from *change*.

We are currently developing the English Vocabulary Profile for levels C1 and C2. This entails looking again at the less frequent meanings of words that are already included in the A1 to B2 levels, and adding new words and phrases.

Accessing the English Vocabulary Profile

You can see a three-letter Preview version for the complete English Vocabulary Profile on the EP website. The current online resource covers A1-B2 levels; click on the thumbnail on the homepage to access the EVP at **www.englishprofile.org**

You can use the English Vocabulary Profile to:

- check the level of words, phrases and meanings
- produce lists of words at certain levels for a particular topic, or with certain grammatical features, etc.
- see how the different meanings of words fit across CEFR levels
- get real examples of how words and phrases are used – by native speakers and by learners at different levels
- compare American English and British English
- get more information on collocations
- get an overview of word families.

For further information about how to use the EVP, visit the EP website and explore the *Information Booklet* under Resources. To find out more about the compilation of the EVP see Capel (2010a, 2010b).

5 How to use the English Profile

This section suggests ways in which ELT professionals such as teachers, curriculum planners and materials or test writers can use the English Profile resources contained in this booklet, specifically to enable them to make decisions about which English language points are suitable for teaching, learning and assessing at each CEFR level. There are four areas listed below which can benefit from the English Profile, with exemplification of how different groups of ELT professionals might use English Profile resources within these areas.

A   **Deciding whether particular English language points are relevant for a specific purpose, learner group and CEFR level**
- A teacher checking whether some key vocabulary for a lesson is suitable for their class.
- A test developer checking whether a particular grammatical point is suitable for an A2 test.
- An author checking what aspects of a grammatical area (e.g. past tense) are suitable for a B1 course.

B   **Identifying suitable English language points for a specific purpose, learner group and CEFR level**
- A curriculum planner is drawing up the vocabulary list for an A1 course.
- An author wants to identify language points that are particularly difficult for Spanish speakers at B1 level.
- A test developer has to decide which structures to include in the assessment syllabus for a C1 exam.
- An author is producing an exercise of prefixes/suffixes for a particular set of words.
- A teacher is looking for a range of examples of 'refusing a request' suitable for B2 learners.

C   **Obtaining authentic learner language to illustrate language points at a specific CEFR level**
- A teacher is putting together an exercise on a particular language point, using examples produced by learners at the same level as their class.
- A test writer is looking for a suitable sentence for a particular test item.
- A curriculum planner wants to add to the syllabus examples of particular structures that are suitable for the level.
- An author is writing a unit on health at B1 level and wants a list of suitable words and phrases to include.
- A teacher is looking for examples of 'asking for permission' in a formal work context suitable for a B2 class.

D   **Gaining a deeper understanding of English language points within and across CEFR levels**
- An author wants to know how an understanding of countable/uncountable nouns progresses from A1 to B1 CEFR levels to work out what should be included in an A1 or B1 level course.
- A teacher wants to see how the different meanings of *keep* are normally acquired across the CEFR levels. It is in the top 500 words for English, but which meanings should students learn first?
- A test writer needs to know what verbs are most suitable for an item on the passive voice at B2 level.
- A curriculum planner wants to make sure the C2 curriculum covers the language of 'presenting a counter-example' in both formal and informal contexts.

The English Profile describes what learners know and can do at each CEFR level. ELT professionals will make use of that information in different ways according to their situation and requirements as there is no single English Profile curriculum that is right for all learners in all contexts. To help ELT professionals, the English Profile will share some sample curricula based on English Profile tools, but ELT professionals will always need to make decisions about how these examples apply to their own context. We welcome your feedback on whether the resources presented here are of potential or actual use in your specific context, together with any examples of how you have already used, or plan to use them.

6 What is English Profile based on?

The English Profile Programme has set out to provide the definitive guide to what learners of English know at each CEFR level. It has done this by combining:

Leading educational organisations	• Cambridge ESOL • Cambridge University Press • The British Council • English UK
World-leading research institutions	• University of Cambridge • University of Bedfordshire • University of Nottingham • and other EP Network Partners
Extensive data about real English language use	• Cambridge English Corpus • Cambridge Learner Corpus • A range of other corpora from around the world • A wide range of coursebooks • Exam specifications and wordlists

English Profile researchers use an innovative methodology for describing English which is both **empirical** in that it is based on real language as produced by speakers of English worldwide; and **international** in that it will not solely be concerned with English as it is spoken in the UK, or in other English-speaking countries. They can achieve this largely because they have access to unique and valuable databases of English language in use.

Previous attempts to describe English at different CEFR levels have been produced by language specialists largely using their insight as expert users and teachers of the language. However, English Profile's results are based on observed data, providing concrete evidence of what learners throughout the world *can do* at each CEFR level. Two of the main corpora used to date are the Cambridge English Corpus (formerly known as the Cambridge International Corpus) and the Cambridge Learner Corpus. The Cambridge English Corpus contains over 1.2 billion words and comprises materials from a huge variety of sources from the UK the USA, Australia and Singapore among others that include real and up-to-date examples of native speaker language. Researchers use this corpus (and others like it) to discover how native speakers use English. For English Profile we also want to look at how non-native speakers use English.

The Cambridge Learner Corpus has been at the centre of EP research work to date (see Section 2 for an introduction), although as it consists entirely of exam data it required complementary non examination learner data. Sometimes learners may try to avoid language they find difficult and stick to "safe" vocabulary, grammar or functions in exams, rather than risk trying out language they're unsure of. This, among other reasons, is why we are currently collecting *non* exam data from learners all over the world for the Cambridge English Profile Corpus (CEPC). We are working to build the Cambridge English Profile Corpus with the collaboration of a network of educational establishments across the world, including state schools, universities, private language schools, research centres, government bodies (including ministries of education) and other ELT professionals. Together we're collecting a wide variety of data types from a number of contexts, such as classroom work, conversations, homework, and so on. The corpus will be balanced across a number of variables, including the first language of learners, the country where data are collected, the age of learners and their CEFR level.

For current information on our corpus-related activities, visit the EP website and click on the *Corpus* page.

7 How to get involved with English Profile

Getting involved with English Profile is your opportunity to shape the CEFR for English. You can submit data for our research, get involved in research itself, or you can simply join the network community and keep in touch with developments. As this publication goes to press we are preparing for the publication of the inaugural volumes of the English Profile Studies series (Green 2011; Hawkins & Filipović 2011) and following that, we will be writing an *English Profile Handbook* full of practical applications, anticipated publication date early 2012. On the events side, EP team members will be presenting at major events in China, Australia and the Czech republic in the coming months (amongst others), do check out our *Events* webpage for more details and we hope to see you there.

Join the EP Network

Joining the EP Network as a data contributor or researcher is straightforward; to submit data please contact us through our website by filling in the form 'Get involved in data collection' under *Community*. Learners submit their data via an online data collection portal which has been developed especially for English Profile.

The benefits of joining the EP Network:

1. Online access to the CEPC (which would include the contributor's own data) in a searchable format. Teachers can use this to help them understand their students' needs better, and to develop teaching materials which cater to those needs.
2. Free access to the English Vocabulary Profile, a fantastic online interactive vocabulary resource developed as part of English Profile.
3. Free tickets to English Profile workshops, which will include training relevant to teachers, such as how to rate a student's work by CEFR level.
4. Advance notice of English Profile-related publications and automatic subscription to *English Profile Journal*.
5. A 'certificate of participation' (on request) for your school, and listing of your school's name, with thanks, on our data contributors webpage.
6. Invitations to English Profile research seminars and other English Profile events.

Give feedback on this booklet

We welcome your feedback on this publication:

- Does it provide the information you need?
- Does the presentation of the information work for you?
- How do you think you could make use of English Profile?
- What else would you like to see from English Profile?

Please go to **www.englishprofile.org** to send us your feedback, mentioning English Profile Version 1.1 .

8 References

Here you will find key references mentioned in this publication. A fuller bibliography can be found on the English Profile website under *Researchers*: **www.englishprofile.org**

English Grammar Profile

Hawkins, J A & Buttery, P (2010) Criterial Features in Learner Corpora: Theory and Illustrations. *English Profile Journal 1*, e5 doi:10.1017/S2041536210000103

Hawkins, J A & Buttery, P (2009) Using learner language from corpora to profile levels of proficiency: Insights from the English Profile Programme. In L Taylor & C J Weir (Eds) *Language Testing Matters: Investigating the wider social and educational impact of assessment*, Studies in Language Testing, vol. 31, (pp. 158-175). Cambridge: UCLES/Cambridge University Press.

Hawkins, J A & Filipović, L (2011) *Criterial features in L2 English: Specifying the Reference Levels of the Common European Framework.* English Profile Studies, vol. 1, Cambridge: UCLES/Cambridge University Press.

O'Keeffe, A & Mark, G (in preparation) *The English Grammar Profile*. To be available online at: **www.englishprofile.org**

Salamoura, A & Saville, N (2010) Exemplifying the CEFR: Criterial features of written learner English from the English Profile Programme. In I Bartning, M Maisa & I Vedder (Eds) *Communicative proficiency and linguistic development: Intersections between SLA and language testing research*, Eurosla Monographs Series, vol. 1, (pp. 101-132). Available online at: **http://eurosla.org/monographs/EM01/101-132Salamoura_Saville.pdf**

Salamoura, A & Saville, N (2009) Criterial features across the CEFR levels: Evidence from the English Profile Programme. *Research Notes 37*, 34-40, Cambridge: Cambridge ESOL. Available online at: **www.CambridgeESOL.org/rs_notes/rs_nts37.pdf**

Saville, N & Milanovic, M (2011) Series Editors' Note in Hawkins, J A & Filipović, L *Criterial features in L2 English: Specifying the Reference Levels of the Common European Framework.* English Profile Studies, vol. 1, Cambridge: UCLES/Cambridge University Press.

English Functions Profile

Online survey: **www.beds.ac.uk/research/bmri/crella/cando**

Green, A (2011) *Language functions revisited: Theoretical and empirical bases for language construct definition across the ability range.* English Profile Studies, vol. 2, Cambridge: UCLES/Cambridge University Press.

Green, A (2010) Requirements for Reference Level Descriptions for English. *English Profile Journal 1*, e6 doi:10.1017/S204153621000005X

Lenz, P & Schneider, G (2004*) Introduction to the bank of descriptors for self-assessment in European Language Portfolios.* Available online at: **www.coe.int/T/DG4/Portfolio/?L=E&M=/documents_intro/Data_bank_descriptors.html**

Saville, N & Milanovic, M (2011) Series Editors' Note in Green, A *Language functions revisited: Theoretical and*

empirical bases for language construct definition across the ability range. English Profile Studies, vol. 2, Cambridge: UCLES/Cambridge University Press.

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

van Ek, J A (1975) *The Threshold Level in a European Unit/Credit System for Modern Language Learning by Adults.* Strasbourg: Council of Europe.

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

van Ek, J A & Trim, J L M (1998a) *Waystage 1990* (revised and corrected edition). Cambridge: Cambridge University Press.

van Ek, J A & Trim, J L M (1998b) *Threshold 1990* (revised and corrected edition). Cambridge: Cambridge University Press.

Wilkins, D (1976) *Notional Syllabuses*. Oxford: Oxford University Press.

English Vocabulary Profile

Cambridge ESOL examination vocabulary lists (KET and PET):
www.teachers.CambridgeESOL.org/ts/digitalAssets/113295_ket_vocablist09.pdf
www.teachers.CambridgeESOL.org/ts/digitalAssets/113298_pet_vocablist09.pdf

Capel, A (2011) *The English Vocabulary Profile*. Available online at: **www.englishprofile.org**

Capel, A (2010a) Insights and issues arising from English Profile Wordlists project. *Research Notes 41*, 2-7, Cambridge: Cambridge ESOL. Available online at: **www.CambridgeESOL.org/rs_notes/offprints/pdfs/RN41p2-7.pdf**

Capel, A (2010b) A1–B2 vocabulary: insights and issues arising from the English Profile Wordlists project. *English Profile Journal*, 1, e3 doi:10.1017/S2041536210000048

Hindmarsh, R (1980) *Cambridge English Lexicon: a graded word list for materials writers and course designers.* Cambridge: Cambridge University Press.

Corpus Development

Alexopoulou, T (2008) Building new corpora for English Profile. *Research Notes 33*, 15-19, Cambridge: Cambridge ESOL. Available online at: **www.CambridgeESOL.org/rs_notes/offprints/pdfs/RN33p15-19.pdf**

Briscoe, E, Carroll, J, & Watson, R (2006) The Second Release of the RASP System. In *Proceedings of the COLING/ACL 2006 Interactive Presentation Sessions*, Sydney, Australia.

Nicholls, D (2003) The Cambridge Learner Corpus - error coding and analysis for lexicography and ELT. In Archer, D, Rayson, P, Wilson, A & McEnery, T (Eds), *Proceedings of the Corpus Linguistics 2003 Conference*, UCREL technical paper number 16, UCREL, Lancaster University.

General English Profile references

Coste, D (2007) Contextualising uses of the common European framework of reference for languages. Paper presented at *Council of Europe Policy Forum on use of the CEFR, Strasbourg 2007*. Available online **www.coe.int/T/DG4/Linguistic/Source/SourceForum07/D-Coste_Contextualise_EN.doc**

Council of Europe (2005) *Reference Level Descriptions for National and Regional Languages. Guide for the production of RLD*. Version 2. Strasbourg: Language Policy Division. Available online at: **ww.coe.int/t/dg4/linguistic/Source/DNR_Guide_EN.pdf**

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

English Profile Bibliography **www.englishprofile.org**

English Profile Glossary **www.englishprofile.org**

Little, D (2007) The Common European Framework of Reference for Languages: Perspectives on the Making of Supranational Language Education Policy. *The Modern Language Journal* 91 (4): 645-55.

McCarthy, M (2010) Spoken fluency revisited. *English Profile Journal 1*, e4 doi:10.1017/S2041536210000012

Milanovic, M (2009) Cambridge ESOL and the CEFR, *Research Notes 37*, 2-5, Cambridge: Cambridge ESOL. Available online at: **www.CambridgeESOL.org/rs_notes/offprints/pdfs/RN37p2-5.pdf**

Saville, N & Hawkey, R (2010) The English Profile Programme – the first three years. *English Profile Journal 1*, e7 doi:10.1017/S2041536210000061

Trim, J L M (2010) The Modern Languages Programme of the Council of Europe as a background to the English Profile Programme. *English Profile Journal 1*, e2 doi:10.1017/S2041536210000097

9 The English Profile Network

Founding Partners
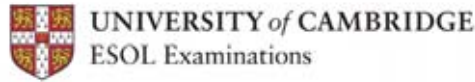
EP Network Partners

The following institutions are involved in data collection or research activities for English Profile:

- ATL Coop Lombardia, Cantù

- Babeş-Bolyai University **www.ubbcluj.ro**

- Banat University of Agricultural Science and Veterinary Medicine **www.usab-tm.ro**

- Colegio Newlands, Buenos Aires **www.colegionewlands.com**

- Cologne University of Applied Sciences **www.fh-koeln.de**

- El Colegio Leonardo Da Vinci, Madrid **www.colegio-leonardodavinci.es**

- Escuela Oficial de Idiomas de Burgos **www.eoiburgos.es**

- Escuela Oficial de Idiomas de Madrid-Moratalaz (Extensión Puente de Vallecas)

    www.educa.madrid.org/web/eoi.moratalaz.madrid/vallecask.html

- FON University **www.fon.edu.mk**

- Hogeschool-Universiteit Brussel **www.hubrussel.be**

- International Programmes Center "EF Penza" **www.efpenza.ru**

- International University of Novi Pazar **www.uninp.edu.rs**

- Kaplan International Colleges **www.kic.org.uk**

- Kragujevac University **www.ekfak.kg.ac.yu** and **www.filum.kg.ac.rs**

- Mary Immaculate College **www.mic.ul.ie**

- Masaryk University **www.muni.cz**

- Megatrend University **www.megatrend.edu.rs**

- Metropolitan University Belgrade **www.fit.edu.yu**

- Moscow Institute of Foreign Languages **www.gaudeamus.ru**

- Novi Sad University **www.uns.ac.rs**

- Parhamergymnasium **www.parhamer.at**

- Palacký University **www.upol.cz**

- Prešov University **www.unipo.sk**

- The Romanian Ministry of Education Research and Innovation **www.edu.ro**

- Shannon College of Hotel Management **www.shannoncollege.com**

- Singidunum University **www.singidunum.ac.yu**

- Talk Tefl Support Centre, Gorgan, Iran

- Technical College Čačak **www.vstss.com**

- Tokyo University of Foreign Studies **www.tufs.ac.jp**

- Universidad Chileno-Británica de Cultura **www.ubritanica.cl**

- Universidad Complutense de Madrid **www.ucm.es**

- Universidad Politècnica de València (Campus de Alcoy) **www.upv.es**

- Universitatea Babeş-Bolyai **www.ubbcluj.ro**

- University of Kragujevac **www.kg.ac.rs**

- University of Niš **www.ni.ac.rs**

- Vilnius Pedagogical University **www.vpu.lt**

- Wyższa Szkoła Lingwistyczna **www.wsl.edu.pl**

EU-funded EP Network Project Members

- Cambridge University Press **www.cambridge.org/uk**

- Cambridge ESOL **www.CambridgeESOL.org**

- CRELLA, University of Bedfordshire **www.beds.ac.uk/research/bmri/crella**

- Banat University **www.usab-tm.ro**

- Cologne University of Applied Sciences **www.international-office.fh-koeln.de**

- Complutense University, Madrid **www.ucm.es**

- Masaryk University **www.muni.cz**

- Presov University **www.unipo.sk**

- The Dublin Institute of Technology **www.dit.ie**

EnglishProfile

Introducing the CEFR for English

The English Profile Programme is an elaboration of the reference level descriptions of the Common European Framework of Reference for Languages (CEFR) that is concerned specifically with the English language.

Since its publication, the CEFR has become influential in building a shared understanding of performance levels for foreign language learners. However, there is a considerable gap between the broad descriptions of levels provided in the CEFR, which cover a range of languages and learning contexts, and the level of detail required for applications such as syllabus or test design, which English Profile addresses.

English Profile has applications in English language pedagogy and assessment, curriculum design, materials development and test construction.

This booklet introduces the theoretical and empirical bases of English Profile, setting out the ambitions of the Programme and presenting emerging findings. It is for teachers, curriculum planners, writers, test developers and other ELT professionals. It will help them make decisions about which English language points are suitable for learning at each level of the CEFR.

Education and Culture DG
Lifelong Learning Programme

COUNCIL   CONSEIL
OF EUROPE   DE L'EUROPE

UNIVERSITY OF
CAMBRIDGE