1 Introduction

1.1 Background: The Common European Framework of Reference (CEFR)

The Common European Framework of Reference (CEFR) has had a major impact on the learning, teaching and assessment of foreign languages in Europe. Originally conceived and developed by the Council of Europe the goals and educational options laid out by the CEFR are set out in the Council's 2001 document, *Common European Framework of Reference for Languages: Learning, Teaching, Assessment* (Cambridge University Press). This document describes how a common framework of proficiency scaling came into being, within the context of the Council's larger cultural and educational goals. What is important as a background to the present book is the Council's proposal that it was possible, and useful, to define six levels of proficiency in the learning of different foreign languages. These levels were given the labels shown in (1):

(1) The CEFR Levels: C2 Mastery
C1 Effective Operational Proficiency
B2 Vantage
B1 Threshold
A2 Waystage
A1 Breakthrough

Whenever different stages of learning and attainment are proposed, one needs some way of distinguishing them. The 2001 publication does this primarily in functional terms, i.e. in terms of the different uses to which language can be put and the various functions that learners can perform as they gradually master a second language (L2). Chapter 3 provides a large number of 'illustrative descriptors' for this purpose. At A2, for example, learners can 'understand sentences and frequently used expressions related to areas of most immediate relevance'. At B1 they can 'understand the main points of clear standard input on familiar matters regularly encountered in work, school, leisure, etc.'. At B2 they can 'understand the main ideas of complex text on both concrete and abstract topics, including technical discussions in his/her field of specialisation'. At C1 learners can 'understand a wide range of demanding, longer texts and recognise implicit meaning', and so on. These examples are all taken from the 'global scale' of Common Reference Levels (Council of Europe 2001:24), which is reproduced here as Table 1.1:

Table 1.1 Common Reference Levels: Global scale

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 and 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.

Many more detailed descriptors are given in the remainder of Chapter 3 for the different levels and with respect to the five skills of listening, reading, spoken interaction, spoken production and writing. The Council of Europe's intention was to reflect social practices in organising learning and to provide a coherent and transparent framework to enable better practices to emerge. In Appendix D (pp.244–257) the authors also reproduce a set of *Can Do statements* developed by the Association of Language Testers in Europe (ALTE), which were anchored to the descriptors and aligned with the CEFR levels in (1). For example, learners at B1 can 'express opinions on abstract/cultural matters in a limited way or offer advice within a known area'. Learners at B2 'can follow or give a talk on a familiar topic'. And learners at C1 'can contribute effectively to meetings and seminars within own area of work or keep up a casual conversation with a good deal of fluency'.

The illustrative descriptors of the CEFR do not give language-specific details about the grammar and lexis that are characteristic of each proficiency level for each L2. Chapter 5 of the 2001 document on 'The user/learner's competences', for example, which includes discussion of syntax, morpho-syntax and lexis, does not link particular grammatical and lexical properties to the CEFR levels with any degree of specificity. There was a reason for this, however: the authors wanted the CEFR to be neutral with respect to the L2 being acquired and to be compatible with the different languages of Europe. In this way a given level of proficiency in L2 German could be compared with a corresponding level in L2 French or English.

The result of this language neutrality, however, is that the CEFR levels are 'underspecified' with respect to key properties that teachers and examiners look for when they assign learners and candidates to a particular proficiency level and score in a particular L2 (see Milanovic 2009). Learners who perform each of the functions in the illustrative descriptors may be using a wide variety of grammatical constructions and words, and the ability to 'do' the task does not tell us with precision how a learner does it and with what grammatical and lexical properties of English (or of other target languages). It is this (deliberate) underspecification that provides the rationale for this book and that explains its subtitle.

The project described in this book is embedded within a larger applied and theoretical research programme, the English Profile Programme (EPP), which was initiated by the Cambridge ESOL group of Cambridge Assessment in collaboration with Cambridge University Press and other stakeholders in 2005. One of the goals of the EPP from the outset has been to provide 'reference level descriptions' and to add grammatical and lexical details of English to CEFR's functional characterisation of the different levels by using the resources of the Cambridge Learner Corpus (CLC). At the time of going to press this is a corpus of roughly 45 million words of written English from learners around the world at all levels of proficiency. The EPP also builds on the pioneering work of van Ek and Trim, for example in their *Threshold 1990* book, which linked many grammatical and lexical details of English to a rich inventory of language functions and notions in clear and practically useful ways. Van Ek and Trim did not have access to the rich electronic resource of the CLC for the empirical testing of their proposals, however, and nor was their work guided by the search for criterial features at the different levels, which is the aim of this book.

1.2 Specifying the reference levels

The basic idea behind the criterial feature concept is that in addition to whether a learner fulfils the communicative functions required by the task or not, there are certain linguistic properties that are characteristic and

indicative of L2 proficiency at each level, on the basis of which examiners make their practical assessments. Since there is a large measure of inter-examiner agreement, and since the illustrative descriptors are underspecified with respect to these L2 properties, we need to discover what it is exactly that examiners look for when they assign the scores they do. Cambridge ESOL has carried out extensive research to help and train examiners to make valid assessments (De Velle 2009, Taylor and Galaczi 2011). It is also reasonable to assume that examiners' collective experience and rater training over many years have led to an awareness of the kinds of properties that distinguish levels and scores from one another. The challenge is to discover what these properties are. This is what the criterial feature concept is all about. If we can make the distinguishing properties explicit at the level of grammar and lexis, and ultimately for phonetics and semantics and form–function correspondences as well, then we will have identified a set of linguistic features that will add the necessary specification to CEFR's functional descriptors for each of the levels. This will have considerable practical benefits for teaching/learning, examining and publishing. It can also contribute new patterns and insights to theories of second language acquisition (SLA).

Milanovic (2009:5) summarises the need for this additional specification as follows:

> The CEFR is neutral with respect to language and, as the common framework, must by necessity be underspecified for all languages. This means that specialists in the teaching or assessment of a given language . . . need to determine the linguistic features which increasing proficiency in the language entails . . . Such features are peculiar to each language and so the CEFR must be adapted to accommodate the language in question. . . . A major objective of English Profile is to analyse learner language to throw more light on what learners of English *can* and *can't* do at different CEFR levels, and to assess *how well* they perform using the linguistic exponents of the language at their disposal (i.e. using the grammar and lexis of English).

Putting this another way, we need to know for each European language and for each level which grammatical constructions are used, which words, which syntactic and morpho-syntactic rules are applied and with what levels of success, and which meanings are assigned to individual words and sentences. And the basic reason why this is vital is because knowing a language and being a competent native speaker means that one has acquired thousands and thousands of properties of English, or Spanish, or French, including the following:

- the sounds of the language
- meaningful units or morphemes
- words (e.g. the nouns and the verbs)

- precise meanings and usage possibilities of words
- basic grammatical constructions
- productive syntactic and morpho-syntactic rules
- exceptions to these, e.g. lexical idiosyncrasies.

As learners progress, they master more and more of these properties, and move closer to the native speaker's knowledge. Through experience and instruction examiners have learned to recognise this progression and to assign examination scripts to the appropriate level. We shall not delve into the thorny issue of what 'native speaker' can mean in SLA. We simply view it as an ideal towards which L2 learners can aspire even though this ideal may be unattainable for many native speakers, who acquired L1 from birth. For an in-depth view and analysis see Davies (1991).

The Council of Europe is to be credited with reorienting language teaching and assessment away from the 'structure-dominated scholastic sterility' (van Ek and Trim 1991:1) that was inherited ultimately from the teaching of classical languages (see E W Hawkins' 1981 *Modern Languages in the Curriculum*, CUP, for a historical summary of language teaching practices) and 'into a vital medium for the freer movement of people and ideas' with its new emphasis on language use and language functions (van Ek and Trim: ibid.). This functional approach within CEFR and in applied linguistics generally can be traced back ultimately to the highly influential work of John Austin in the philosophy of language, as captured most explicitly in his 1962 book *How to Do Things with Words*. This book was a reaction against a long tradition of research in logical semantics focusing on basic sentence types that carry descriptive or 'truth-conditional' meanings, in favour of a new usage-based and 'speech act' approach (to use John Searle's term which has now largely replaced Austin's 'performative', see Searle 1969). The usage-based philosophy of language provided a theory and a vocabulary for describing the many uses to which language can be put and the manner in which the functions are expressed, whether through 'direct' or 'indirect' speech acts, etc. (see Searle 1975). The influence of this theoretical work on the CEFR, and on the detailed functional taxonomy of van Ek and Trim's *Threshold 1990*, is unmistakable. See Anthony Green (forthcoming 2012: Chapter 2) for a detailed literature review and a historical survey of the origins of the CEFR's language functions going back to Austin, Searle and other philosophers of language.

One reason why we now need to return to a greater focus on grammar and lexis, as part of this added specificity for the different levels in different languages, is because there is, in fact, no simple one-to-one correspondence between functions and linguistic forms. One and the same sentence type can perform many functions. *Can you take out the garbage?* can be an (indirect) request, and also a (direct) question about your ability to do something.

Conversely, one and the same function (requesting or commanding) can be performed by many linguistic structures or single words. Most of the basic construction types of a language, for example intransitive versus transitive versus ditransitive clauses, and most single words and phrases have syntactic and semantic properties that are independent of, though compatible with, a whole range of functions that can be performed using them, and it would complicate matters to make reference to these functions when describing their basic grammar and semantics. Similarly, the types of grammatical errors that learners make at different levels, in inflection, derivation, agreement, word order, etc., are not in general aligned with the kinds of functions they are trying to perform at these levels. There are some clear form–function correspondences, of course, and it is practically useful for learners to have them listed and pointed out. But any functional approach that contains a list of the functions that learners can express at given levels, whether with their common grammatical and lexical exponents or not, needs to be supplemented by a description of the partly orthogonal and autonomous syntactic, morpho-syntactic and lexical properties of the language that are characteristic of the different levels. To quote from Milanovic (2009:5) again:

> We are now in a position to begin 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.

In order to realise this, we need better descriptions of what second language learners actually know as their learning progresses. Let us divide the learning process into six or eight or four stages, or however many the learning data enable us readily to discriminate. The CEFR proposes six, as we have seen, and examining boards have operated with six, prior to and independently of the CEFR. In other words, six levels have been widely regarded as useful. So what are the characteristic properties for a language like English of these six learner levels, in phonetics/phonology, morphology, syntax, the lexicon, semantics, pragmatics and discourse? What do we expect to find at level two or four or six across each of these areas? What phonological skills, lexical-semantic knowledge, morpho-syntactic error types and syntactic patterns correlate at each of the levels? There is, we submit, no theory of SLA that can successfully predict these correlations across the broad range of language mastery skills for each of six levels. Yet this is exactly what we need to specify if learners, teachers, teacher trainers, examiners, curriculum developers and publishers are to do their respective jobs better, guiding learners and teachers more effectively to the next stage, publishing materials that are

better oriented to learners' needs, making better assessments, and also training examiners better.

Practitioners have, of course, accumulated considerable knowledge, sometimes implicit sometimes explicit, about these learning stages and their specific properties. Examiners have learned, as suggested above, to agree with one another on the appropriate level to assign to an arbitrary script, based on their training and on years of practical experience assigning scores and passes or fails to exams at various levels. Along with the realisation of specific functions that the task requires, one aspect of performance that they look out for when they make these practical assessments is the details of the grammar, lexis and semantics of English. These linguistic properties can, we submit, often be independent of the functional descriptions of the CEFR, even when the examination level is set within the six-level Common European Framework. The validation systems and statistical tests that support these assessments, certainly for exams administered by Cambridge ESOL, have always been based on the properties of learner English at the different levels and on empirically derived scales and psychometric tests, in addition to the functional descriptors themselves (see Milanovic 2009). In short, as already indicated, we need to specify the reference levels of the Common European Framework for different languages.

Using the Cambridge Learner Corpus

We are fortunate to have at our disposal a rich empirical resource for this purpose, the Cambridge Learner Corpus (CLC), which has been developed over many years by Cambridge University Press and Cambridge ESOL. Roughly half of the CLC's current 45 million words of written learner data is coded for errors. Full details on the CLC and the type of data it contains are given in Chapter 3. The CLC was originally searchable lexically, i.e. on the basis of individual words, and grammatically only to the extent that a rule of English grammar was reflected in an error code. This search capability has now been expanded and the CLC has been tagged for parts of speech and parsed by Ted Briscoe of the Cambridge Computer Lab and Paula Buttery of RCEAL Cambridge using the Robust Accurate Statistical Parser (RASP) developed by Ted Briscoe and John Carroll (see Briscoe, Carroll and Watson (2006) accessible at: http://acl.ldc.upenn.edu/P/P06/P06-4020.pdf). RASP is an automatic parsing system incorporating both grammatical information and statistical patterns, and details of its operation are summarised in Chapter 3. The CLC therefore provides empirical patterns of language use that can inform our search for the specifics of each learner level.

The CLC's error codes have been developed by computational linguists at Cambridge University Press. They classify some 76 error types involving lexical, syntactic and morpho-syntactic properties of English. A small sample

is given in (2) together with exemplifying sentences. Note that none of these sentences or other sentences that illustrate error codes in this book are taken from corpus data, but are provided for illustrative purposes only:

(2) Sample Error Codes in the CLC

RN	**Replace Noun**	*Have a good travel* (journey)
RV	**Replace Verb**	*I existed last weekend in London* (spent)
MD	**Missing Determiner**	*I spoke to President* (the) *I have car* (a)
AGV	**Verb Agreement Error**	*The three birds is singing* (are)
IV	**Incorrect Verb Inflection**	*I spended last week in London* (spent)
FJ	**Wrong Adjective Form**	*The situation got worst* (worse)
UQ	**Unnecessary Quantifier**	*A little bit quite common* (quite common)
DY	**Derivation of Adverb**	*It happened fastly* (fast)

The CLC also contains data from numerous (over 130) typologically and genetically different first languages.

One of the strengths of an empirically based corpus approach to learning such as this is that we can focus not just on errors (i.e. on what learners get WRONG), but on what they get RIGHT. Using the corpus we can quantify, for each learning stage, how many of the thousands of properties that constitute knowledge of English learners actually use. We can also measure how their linguistic performance gradually improves relative to that of native English speakers. In order to compare the learner data with actual English usage by native speakers we can search the British National Corpus (BNC). The BNC comprises 100 million words of modern British English, from a wide range of sources and text types (90 million written, 10 million spoken). It has been tagged and parsed using the same automatic parsing system (RASP) that has been applied to the CLC, making exact comparison between the CLC and the BNC possible.

The CLC gives information for each script on whether the candidate passed or failed and on the candidate's score (A to F). It is important to stress that for the purpose of this research we examined only those scripts with passing grades of A, B and C and filtered the fails. The reason for this is a principled one. We are focusing here on what passing candidates have actually learned, i.e. we are trying to define the criterial features of pass scripts at each CEFR level that learners need to master in order to satisfy the requirements for success at that level. We are aware that scripts that have been graded by examiners as being below the satisfactory band on the mark scheme do also contain correct uses of many syntactic frames and lexical items, but some of these scripts must have been marked as unsatisfactory because fewer correct structures and items were used and more errors were made in general.

We are also fortunate, in our study of learner English, to have the benefit, as stated above, of previous work by van Ek and Trim, specifying the English language functions and notions that can be expressed by learners at different levels together with their suggested grammatical and lexical exponents. Details of the Threshold level for English were set out in van Ek and Trim's *Threshold 1990* published by Cambridge University Press in 1991. *Waystage 1990* was published by van Ek and Trim in the same year, *Vantage* in 2001. The volume *Breakthrough*, specifying CEFR A1 English language level competences, has recently been made available electronically for EPP by permission of John Trim and is accessible via the English Profile website (www.englishprofile.org).

1.3 What are criterial features?

The title of this book promises 'criterial features'. We need to clarify what these are, and exemplify them, at the outset. The basic idea is that we try to find properties of learner English that are characteristic and indicative of L2 proficiency at each of the levels and that distinguish higher levels from lower levels.

For example, Caroline Williams (2007) has identified when the basic construction types of English first appear in the CLC. Simple intransitives (NP-V) and the slightly more complex transitive (NP-V-NP) sentence types are present from the beginning at the **A** levels (Williams examined only data from A2 onwards, but they are there at A1 as well):

A1 *He went.* (NP-V)

A1 *He loved her.* (NP-V-NP)

Modal auxiliary verbs like *may*, *might*, *can* and *must* appear first at **A1** or **A2**, but only in some of their senses. See the English Vocabulary Profile (EVP), previously known and referred to in this volume as English Profile Wordlists (Capel 2010), currently available in preview form via the EP website (www.englishprofile.org). *May* is first attested at **A2** in its epistemic sense of POSSIBILITY, as in the following attested example:

A2 *Then we may go sightseeing*. (POSSIBILITY)

but not yet in its deontic PERMISSION sense (which is **B1**, see below). *Can*, on the other hand, is first attested in the PERMISSION sense at **A1** and in the POSSIBILITY sense at **A2**, as in the following attested examples:

A1 *And if you want, you can bring pencils or pens*. (PERMISSION)

A2 *It is an interesting place because you can see a lot of plants.* (POSSIBILITY)

Lexical verbs appearing at the **A** levels are typically among the most basic and frequent verbs of English (see Hawkins and Buttery 2009 for exemplification

and a comparative quantification for selected verbs in the CLC and the BNC), and they typically appear first in their most basic and frequent senses. Verbs attested at **A1** include *catch*, *eat*, *give*, *put*, *take* and *walk*. New verbs at **A2** include *break*, *cut*, *fall*, *hit*, *push*, and *stand*, within an expanded total of verbs, again typically in their most basic and literal senses. For *break* this includes its primary physical sense as in the attested:

A2 *I broke a beautiful glass.*

and for *cut* it includes the following attested example of its primary sense:

A2 *First I cut the cake with my mother*.

Properties found at the lower levels of the CLC generally persist through the higher levels. Those that appear first at A2 discriminate only between A1 and all other levels, and their usefulness as criterial features is limited. The new features at B1 are more interesting for criteriality. For example, Williams (2007) found that Object Control structures with an *-ing* verb complement of the type

B1 *I caught him stealing.*(NP-V-NP-V (+*ing*))

appear first at B1 and are criterial for this and for all higher levels, distinguishing them from A1 and A2. Our research has shown that structures with a finite complement clause postposed to the right of predicates like *is true* and *seem* with a subject *it* in so-called 'Extraposition' structures, are also criterial for B1 and higher levels. The following is an attested example:

B1 *It's true* [*that I don't need a ring to make me remember you*]
(*it*-*be*-Adj-S)

The modal auxiliary verb *may* is first used in its deontic sense of PERMISSION at **B1**, as in the following attested example:

B1 *May I borrow your bicycle for the weekend?*

in contrast to its epistemic POSSIBILITY meaning at **A2** above (see the English Vocabulary Profile). A large number of lexical verbs including *divide*, *fit*, *grab*, *spill*, *stick* and *tear*, appear for the first time at **B1** within a further expanded total of verbs. And the meanings of the lexical items that appeared first at A1 and A2 begin to expand from their basic senses above. So *break* appears for the first time in the extended sense of INTERRUPT at **B1**, in the attested example:

B1 *I think the most important aim of a holiday is to break your daily routine.*

Constructions that are criterial for **B2** and the higher levels include 'secondary predications' with object control such as

B2 *He painted the car red.* (NP-V-NP-AdjP, Object Control)

with *red* predicated of the direct object *the car*, as well as another Object Control structure (see Williams 2007):