Exemplifying the CEFR: criterial features of written learner English from the English Profile Programme

Angeliki Salamoura & Nick Saville
University of Cambridge ESOL Examinations

English Profile (EP) is a collaborative programme of interdisciplinary research, whose goal is to provide a set of Reference Level Descriptions (RLDs) for English for all six levels of the Common European Framework of Reference (CEFR). This chapter summarises work and outcomes to date from one of the EP research strands which focuses on corpus linguistics, second language acquisition, psycholinguistics and computational linguistics. The findings discussed are based on the Cambridge Learner Corpus, a database of over 39 million words of written English produced by English learners from around the world, taking Cambridge ESOL examinations. The chapter illustrates how hypotheses formulated from models of second language acquisition (SLA) and psycholinguistics, and a corpus-informed approach are used to investigate second language learner data in order to develop the RLDs for English. By adopting such an approach, English Profile aims to produce an exemplification of the CEFR informed by SLA theory and at the same time, shed light to questions and issues raised by SLA and psycholinguistic theory based on empirical (learner) data. A main focus within EP is the identification of 'criterial features', i.e. features from all aspects of language which can distinguish CEFR levels from one another and thus serve as a basis for the estimation of a learner's proficiency level.

1. Introduction: The English Profile Programme

The English Profile Programme (henceforth EP) is a collaborative programme of interdisciplinary research, whose goal is to provide a set of Reference Level Descriptions (RLDs) for English for all six levels of the Common European Framework of Reference (CEFR) from A1 to C2 (Council of Europe, 2001; see Little, 2007, pp. 167-190 for an extended discussion of the CEFR). (For an overview of the EP research programme see Kurteš and Saville, 2008; Salamoura, 2008.) One strand of the EP research programme focuses on *corpus linguistics, second language acquisition, psycholinguistics* and *computational linguistics*. This chapter will summarise work and outcomes to date from this

EUROSLA MONOGRAPHS SERIES 1
Communicative proficiency and linguistic development, 101-132

research strand.[1] In particular, it will illustrate how hypotheses formulated about prevalent issues in second language acquisition (SLA) and psycholinguistics, and a corpus-informed approach, are used to investigate second language learner data in order to develop RLDs. By adopting such an approach, EP aims to produce an exemplification of the CEFR informed by SLA theory and at the same time shed light on questions and issues raised by SLA and psycholinguistic theory based on empirical (learner) data. The EP aims and approach reflect strongly the objectives and concerns of the SLATE network as detailed in the introductory chapter of this volume (Hulstijn, Alderson, & Schoonen, this volume). This close relationship will become apparent in the remainder of this chapter, where references will be made to SLATE's research questions and goals as presented in the introductory chapter of this volume (see sections "SLATE's overarching research question" and "SLATE's more specific research questions and research goals").

A main focus within the programme is the identification of 'criterial features' for each CEFR level, or in other words, how each level differs from adjacent levels (cf. Hendriks, 2008). This focus closely matches, in fact, SLATE's central research question as formulated in the Introductory chapter, particulary research question 4 (Hulstijn et al., this volume). Of course, EP is concerned with the identification of 'criterial features' for L2 English. 'Criterial features' are linguistic properties from all aspects of language (phonology, morphology, syntax, semantics, discourse, etc.) which can distinguish the different proficiency (here CEFR) levels from one another and thus can serve as a basis for the estimation of a learner's proficiency level. In fact, EP researchers have drawn an analogy between criterial features and the defining characteristics for recognising faces in a police identikit. In an identikit, one does not need to see all the features of a person's face in order to distinguish that person from others; the important defining characteristics that capture essential qualities are typically enough for such a distinction. Criterial features operate in a similar way – they capture essential distinguishing properties of the CEFR proficiency levels (Hawkins, MacCarthy, & Saville, personal communication, April 12, 2010).

What makes a feature 'criterial' is an open question which the EP researchers have been addressing as part of their collaborative research agenda. The programme has adopted an iterative approach to formulating and testing research questions and hypotheses: as empirical evidence is accumulated and

1 For work and outcomes in other more pedagogically and assessment oriented EP research strands, see e.g. Green (2008) and Capel (2009).

shared, more criterial features will be identified. The more the criterial features are understood in relation to the empirical data, so the research questions will be refined over time.

A more comprehensive definition of criterial features is discussed in a following section (*Towards a definition of criterial features*). We will define four types of feature whose use or non-use, accuracy of use or frequency of use may be criterial for distinguishing one CEFR level from the others: (i) acquired/learned language features, (ii) developing language features, (iii) acquired/native-like usage distribution of a correct feature, (iv) developing/non native-like usage distribution of a correct feature. We will also provide examples of linguistic features that have been identified as criterial from the hypotheses formulated and tested thus far.

Although informed by SLA, EP research does not aim to put to the test existing SLA theories or compare and contrast competing SLA models. As [the authors of the introductory chapter in this volume] demonstrate, there are inherent problems in trying to find direct links between proposed acquisitional stages in SLA theory and any of the levels defined in the functional and formal CEFR scales, as these two may not necessarily coincide and, in fact, they were not designed to coincide – the CEFR authors clearly say that the Framework is deliberately atheoretical (Council of Europe, 2001). Instead, Hulstijn et al., (this volume) argue that a potentially fruitful and meaningful research approach for the exemplification of the CEFR should be directed at trying to characterize successful performance at a given functional CEF level, in terms of *how learners express meaning with linguistic form*. Once linguistic forms and their equivalent functions per CEFR level have been identified, the next step would be to find an *explanation* for these findings. The EP research approach follows a similar pathway.

The source of the findings reported in this chapter is the *Cambridge Learner Corpus* (CLC), a database which currently comprises approximately 39 million words of written production from over 160,000 learners of English from a wide range of L1 backgrounds and, critically, is linked to the CEFR (see the next section for a detailed description of the CLC). Starting from such an extensive empirical database, EP research aims at identifying systematic patterns in the learner data either inductively or deductively, which current theories or models of SLA do not necessarily predict. Informed by current issues in SLA and related disciplines (e.g. frequency of input, L1 transfer, etc.), we then formulate explanatory principles that can account for these emerging data patterns, and in turn, can inform current SLA issues. This approach is explained in Hawkins and Filipović (2010):

> In order to find the criterial features of a level we use a mix of inductive and deductive techniques. The CLC reveals many patterns that are not theoretically predictable and that emerge in response to inductive search queries. We also proceed deductively by searching selectively in the corpus for grammatical and lexical patterns that we believe will be distinctive for the different levels, after consulting a broad range of linguistic and psycholinguistic theories that help us make informed decisions about what is likely to be criterial. These theories come from studies of first and second language acquisition, language processing, grammatical complexity, the lexicon and lexical semantics, and language typology. A set of hypotheses was formulated at the outset of the EPP for emerging patterns in second language acquisition, derived from these theories, which were then gradually tested and refined as the project developed.

As emphasised above, EP is a collaborative, interdisciplinary programme of research involving a number of researchers working on different but ultimately interrelated aspects of the EP research agenda. The EP research work referred to in this chapter has been carried out by or under the supervision of Prof. John Hawkins of the Research Centre for English and Applied Linguistics at the University of Cambridge, and in collaboration with the authors who have been critically involved in the EP research agenda in a variety of roles, including commissioning research, reviewing and interpreting research findings and planning future research directions. This chapter aims to collate the main EP research findings as discussed and presented in a number of interim publications that have, thus far, been circulated as internal reports and papers within the EP research circle or presented at internal EP seminars or meetings.

As mentioned above, the research findings discussed in this chapter are based on analyses of learner data from the Cambridge Learner Corpus, a detailed description of which is provided in the next section.

2. Cambridge Learner Corpus

The Cambridge Learner Corpus (CLC) is a unique collection of learner written English, developed since the early 1990s by Cambridge University Press and the University of Cambridge Local Examinations Syndicate (now Cambridge ESOL). At the time of writing the CLC consists of approximately 39 million words of learner written English produced by candidates taking Cambridge ESOL examinations. Approximately 20 million of these data are error-coded. (For the latest updates on the size and scope of the CLC, please consult the official website at http://www.cambridge.org/fi/elt/catalogue/subject/custom/item3646603/Cambridge-International-Corpus-Cambridge-Learner-Corpus/?site_locale=fi_FI)

Alongside the candidates' written responses to the Writing examination papers (extended writing tasks which require candidates to produce an extended piece of text as opposed to short answer questions or cloze tasks) the corpus contains information about the candidates, such as their gender, age, first language, reason for taking the exam etc., and the candidate's overall mark or grade and marks on the other components (typically Reading, Listening and Speaking). The tasks to which the candidates responded are also available, both as images and as a searchable sub-corpus.

At present, EP is using data from a subset of the CLC, amounting to some 26 million words (half of which are error-coded). The examinations in this corpus subset are Main Suite tests (a general purpose suite) consisting of: Certificate of Proficiency in English (CPE), Certificate in Advanced English (CAE), First Certificate in English (FCE), Preliminary English Test (PET), and Key English Test (KET).

The examinations are aligned with the CEFR (Council of Europe, 2001) as shown in Table 1 (adapted from Taylor, 2004, p. 3; see also Jones, 2000, 2001, 2002; Taylor & Jones, 2006, for empirical validation of the alignment). Thus, the EP findings reported in this paper derive from data ranging from A2 to C2 (CF. Table 1) and A1 level may be not mentioned in analyses.

Table 1: Alignment of Cambridge ESOL Examinations with the CEFR scale.

CEFR level	Descriptive title	Main Suite
A1	Breakthrough	
A2	Waystage	KET
B1	Threshold	PET
B2	Vantage	FCE
C1	Effective Operational Proficiency	CAE
C2	Mastery	CPE

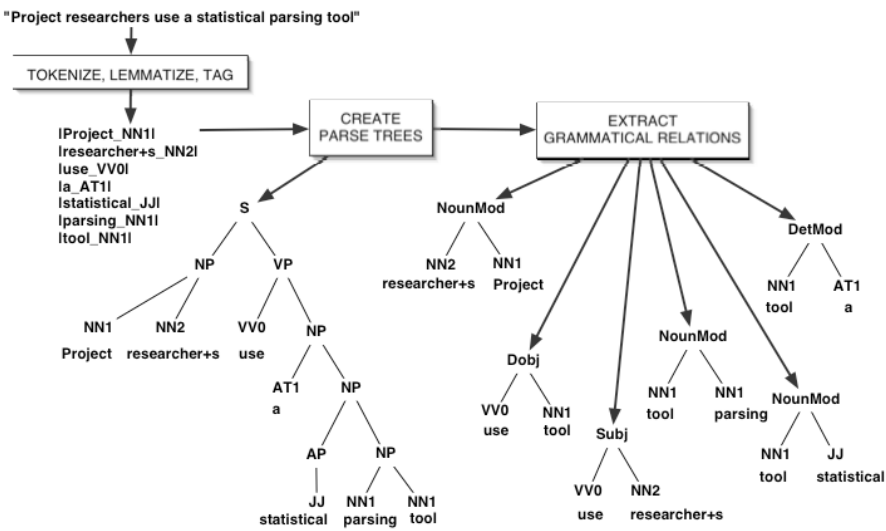
The CLC's system of error codes consists of over 70 codes, each containing two parts: the type of error and the part of speech it applies to. Some examples are provided in Table 2 (after Hawkins & Buttery, 2009a; see Nicholls, 2003, for a description of the error-coding system).

Table 2: Sample error codes in the Cambridge Learner Corpus.

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

The existence of these error codes together with the meta-data about the candidate and the exam enabled the calculation of frequency statistics for each exam level, language group, age, etc. (see Tables 4 and 5 on determiner errors in the criterial features section below). To enable searches beyond the individual word level and over a wide range of lexical and grammatical features, the CLC was subsequently tagged for parts-of-speech and parsed using the Robust Accurate Statistical Parser (RASP) (Briscoe, Carroll, & Watson, 2006). RASP is an automatic parsing system

Figure 1: Parsing of 'Project researchers use a statistical parsing tool' using RASP (from Hawkins & Buttery, 2009a, p. 162).



incorporating both grammatical information and statistical patterns, and its operation is summarised by Hawkins and Buttery (2009a, p. 161).

> When the RASP system is run on raw text, such as the written sentences of the CLC, it first marks sentence boundaries and performs a basic 'tokenisation'. Part-of-speech tags are assigned in a probabilistic basis. The text is then 'lemmatized', based on the tags assigned to word tokens. For each sentence a parse forest representation is generated containing all possible parse trees and subanalyses, with their associated probabilities. And a weighted set of grammatical relations is extracted associated with each parse tree. These operations are shown in Figure 1 (*see next page*) for the sample sentence 'Project researchers use a statistical parsing tool', using just one illustrative parse tree and its associated grammatical relations.

For more details on the annotation of the CLC with part-of-speech tags, word lemmas, grammatical relations and complexity metrics, as well as illustrations see Hawkins and Buttery (2009a, pp. 169–172).

3. CLC and SLA

CLC offers a rare opportunity to explicate the CEFR levels for the English language and at the same time explore a number of SLA issues. Previous research in SLA has been constrained by three main sets of limitations. The first set pertains to the reliable identification of the proficiency level of the study participants. Not many SLA studies provide a systematic or assessment-based classification of learners according to proficiency, or if they do, the measurement tools used vary from study to study.[2] Moreover, the varied use of terminology for level description (which is not always adequately defined), e.g. "advanced learners vs. beginners", "high proficiency vs. low proficiency learners", does not help either. As a consequence, the degree of generalisabilty and comparability of findings across different SLA studies has always been a major caveat in these studies. Furthermore, the variability exhibited in the

2 There are, of course, notable exceptions to this statement (e.g. Bialystok, 2001; Hulstijn, 2006; Kessler, 2007; Pienemann & Kessler, 2007; Skehan, 1998, etc.) who have addressed the issue of defining and measuring second language profiles in a systematic way. Pienemann's (1992) COALA programme also represents an early attempt to systematically and reliably classify stages of SLA from empirical data.

learner interlanguage (IL) makes the systematic classification of learners into levels an all important issue for SLA research. It has been observed, for instance, that learners show more variability than native speakers in terms of target language forms, e.g. learners may alternate between forms (e.g. *no* and *not*) to express the same language function (negation) in a non-systematic way and only with increased proficiency do they establish a one-to-one from/function relationship (Gass & Selinker, 2008, p. 259). Or, learners may alternate between use or nonuse of a form, e.g. plural marking, even on the same lexical item (Young, 1991). Given the existing variability in second language learning, second language data can only be a useful tool for research and pedagogic purposes if their proficiency level has been reliably identified, e.g. via valid alignment to a widely accepted proficiency framework, such as the CEFR. CLC's learner data fulfil this requirement.

In the current strand of EP research, learner proficiency (linguistic ability) at each CEFR level will be described in terms of the following properties (Hawkins & Buttery, 2009a, p. 160):

- meaningful units or morphemes;
- lexical items (e.g. nouns and verbs);
- basic grammatical constructions;
- productive syntactic and morpho-syntactic rules;
- exceptions to some of these, i.e. lexical idiosyncrasies.

We do not argue that this is a comprehensive definition of second language proficiency or ability. It is one that covers the lexico-semantic, morpho-syntactic and syntactic aspects of language learning which is the focus of the EP research described in this paper.

The second set of limitations involves the features studied. Most SLA research to date has focused on the investigation of single language features or phenomena across levels (the vertical or developmental dimension of SLA), thus offering little information about the co-occurrence of different developing SLA features within the same proficiency level or, most importantly, about their interaction. The CLC provides a large empirical database of L2 data and an advanced search capability, as outlined above and detailed in Hawkins and Buttery (2009a). These two features allow EP researchers to search for correlations among a large set of diverse lexical and grammatical features and thus draw conclusions about the interrelation of developing SLA features, i.e. the horizontal principles of SLA mentioned above.

The third set of limitations in SLA research concerns the combination of L1s commonly used in L1 transfer research. The majority of SLA studies reach preliminary conclusions based on the study of a handful of L1s. In contrast, CLC comprises data from learners from 130 L1s, thus permitting an extensive study of L1 transfer effects across all major language families. Cross-linguistic differences per CEFR level is one of the main premises under investigation in EP, reflecting SLATE's objective of examining the extent of the source language (L1) involvement in determining a learner's linguistic profile (research question 2 in Introductory chapter, this volume).

4. Towards a definition of 'criterial features'

As discussed in the Introduction, a main focus of EP is the identification of 'criterial features' of English for each CEFR level, which will differentiate one level from adjacent levels. By criterial features we mean features that characterise and distinguish the six CEFR levels. By definition then the criterial features form a subset of all possible features that may appear at a given level and by definition these criterial features can be used, by virtue of their distinctiveness, for diagnostic purposes in language learning, teaching and assessment. For example, the occurrence or not of criterial features in a learner's output can diagnose their CEFR level and/or distinguish them from other learners whose use of the same criterial features differs (significantly) from that of the first learners.

In the *Corpus and Computational Linguistics* strand, Hawkins and Buttery (2009b) have identified four types of feature that may be criterial for distinguishing one CEFR level from the others. Although couched primarily in grammatical terms (i.e. lexical semantic, morpho-syntactic and syntactic features), this classification may also be extended to encompass other types of language features. The four categories, illustrated with examples, are as follows:

1. Acquired/Learned language features

These are language features a learner masters at a given level and uses accurately and consistently at the higher levels. In this category fall the *'positive grammatical properties'* that Hawkins and Buttery (2009b) describe as:

> …correct properties ... that are acquired at a certain L2 level and that generally persist at all higher levels. E.g. property P acquired at B2 may differentiate [B2, C1 and C2] from [A1, A2 and B1] and will be criterial for the former. Criteriality characterises a set of adjacent levels in this case. Alternatively some property might be attained only at C2 and be unique to this highest level.

For instance, verb co-occurrence frames appearing for the first time at B1 level, e.g. NP-V-NP-NP structures (*She asked him his name*), are criterial for [B1, B2, C1, C2], whereas new verb co-occurrence frames appearing at B2, e.g. NP-V-NP-AdjP (Obj Control) (*He painted the car red*), are criterial for [B2, C1, C2] (see Tables 3b-c).

2. Developing language features

These are features that appear at a certain level but they are unstable, i.e. they are not used correctly in a consistent way. This category includes what Hawkins and Buttery (2009b) call 'negative grammatical properties of an L2 level, i.e.:

> …incorrect properties or errors that occur at a certain level or levels, and with a characteristic frequency. Both the presence versus absence of the errors, and the characteristic frequency of the error (the 'error bandwidth') can be criterial for the given level or levels. E.g. error property P with a characteristic frequency F may be criterial for [B1 and B2]; error property P′ with frequency F′ may be criterial for [C1 and C2].

Hawkins and Buttery (2009b) define criteriality for "*negative grammatical properties"*, i.e. errors, as follows:

An error distribution is criterial for a level L if the frequency of errors at L differs significantly from their frequency at the next higher and lower levels, if any. Significance amounts to a difference of at least 29% from level to level, which guarantees at least one standard deviation from the mean. Two or more levels can be grouped together for criteriality if each is not significantly differentiated from any immediately higher and lower levels (i.e. by less than 29%).

For instance, preposition errors (e.g. *When I arrived at London*) do not show significant differences in frequency of occurrence between B1 and B2 or between B2 and C1, but their frequency of occurrence drops significantly from C1 to C2 level. Therefore, the relevant "error bandwidth", as defined above, is criterial for B1, B2, C1 versus C2. As explained above, the three levels (B1, B2, C1) are grouped together in this case since they are not significantly different from each other with respect to their error frequency/bandwidth.

Given the evolving nature of second language acquisition/learning, one would predict that several language features would pass through a developing stage before they are acquired/learned. So, one feature that is still developing at one proficiency level may be acquired at the next level up, or a feature may be developing across more than one level.

Although we analyse incorrect language properties, we do not conduct a mere error analysis in its traditional form (e.g. Corder, 1967). We are not looking at individual, random errors; we are looking at error patterns; we are not only considering errors, i.e. what learners cannot do (cf. (2) and (4) below), but also what learners can do at the same CEFR level (cf. (1) above and (3) below). It is the combination of correct and incorrect language features in a learner's IL that will provide a comprehensive insight into learners' second language performance and overcome shortcomings inherent in traditional error analysis (see e.g. Gass & Salinker, 2008, p. 102ff). As Gass and Selinker (2008) note, errors "provide windows onto a system – that is, evidence of the state of a learner's knowledge of the L2" (p. 102). It is in this sense that we take into account incorrect properties when defining and identifying criterial features.

3. Acquired/Native-like usage distributions of a correct feature

> Positive usage distributions for a correct property of L2 that match the distribution of native speaking (i.e. L1) users of the L2. The positive usage distribution may be acquired at a certain level and will generally persist at all higher levels and be criterial for the relevant levels, e.g. [C1 and C2] (Hawkins & Buttery, 2009b).

For example, the relative distribution of indirect object/oblique relative clauses (*the professor that I gave the book to*) at C1 (4.63%) and C2 levels (4.29%) in relation to relatives on other positions *(*subjects, direct objects and genitives) in CLC learner data approximates the distribution observed in the British National Corpus (BNC; at least 4.33% - see Table 6 for the usage of different types of relative clauses as percentage of total within each CEFR level in the CLC learner data). This distribution is, therefore, a strong candidate for a criterial feature (acquired usage distribution) for C1 & C2 (Hawkins & Buttery, 2009b).

4. Developing/Non native-like usage distributions of a correct feature

> Negative usage distributions for a correct property of L2 that do not match the distribution of native speaking (i.e. L1) users of the L2. The negative usage distribution may occur at a certain level or levels with a characteristic frequency F and be criterial for the relevant level(s), e.g. [B2] (Hawkins & Buttery, 2009b).

The same distribution described above, i.e. the relative distribution of indirect object/oblique relative clauses (*the professor that I gave the book to*) to relatives in

other positions (subjects, direct objects and genitives), at CEFR levels A2 (1.61%), B1 (1.62%) and B2 (2.80%) in the CLC departs significantly from the typical usage distribution of native speakers (at least 4.33% in BNC - see Table 6 for percentages of relative clause usage in the CLC learner data). This distribution is thus an example of a developing usage distribution at A2-B2 levels and is criterial for A2, B1 and B2 versus C levels (Hawkins & Buttery, 2009b).

Of course, not all criterial features will have diagnostic power at an individual learner output/script level as this was described at the beginning of this section. A number of both acquired and developing properties, defined as criterial for a certain level based on frequency of occurrence across all scripts (or a wide range of learner data), may be absent from an individual script not necessarily because the learner hasn't mastered them but simply because the small size or specific focus of the individual script may not allow or encourage the use of these properties. Or in other words, there is also the question of how dense the learner data is in terms of criterial features – the absence of a sufficient amount of such features may also be due to too little evidence of structures in the individual learner data rather than the unsuitability of criterial features for diagnostic purposes.[3] Hawkins and Filipoviç (2010) are currently investigating this issue further in an attempt to capture the diagnostic relevance of criterial features when applied to individual scripts. Identifying which criterial features can also serve as diagnostics for language learning, teaching and assessment purposes is one of the main research leads pursued within EP and is in accordance with SLATE's concerns about "which linguistic features, emerging from [CEFR] profiling research, can serve as successful tools in the diagnosis of learners' proficiency levels and of weaknesses that require additional attention and training" (research question 4 in Introductory chapter, this volume).

Moreover, the aforementioned four types of criterial features describe not only what learners *can do* (types 1 and 3) but also what learners *cannot do* (type 2) and what learners *cannot do as well as or to the same extent as* native speakers (type 4). All these three aspects of language performance form fundamental parts of second language learning which, according to the SLATE group, should be investigated to identify the limits of learners' performance at each CEFR level (research question 3 in Introductory chapter, this volume).

3 We are grateful to an anonymous reviewer for this remark.

5. Criterial features and SLA

The search for the above four types of criterial features within the CLC is primarily driven by current issues and questions in SLA theory (e.g. frequency of occurrence of L2 structures, L1 transfer; see Doughty & Long, 2005, for an overview of such issues), as well as by psycholinguistic principles of processing efficiency and complexity (Hawkins, 2004). Based on Hawkins' (2004) theory of *Efficiency and Complexity in Grammars*, a number of general patterns and principles of developing SLA stages of English have been identified using the CLC. These principles and patterns, in turn, enable us to define a number of criterial features across the CEFR levels. This section will illustrate three of these principles drawing on Hawkins and Buttery (2009a), and Hawkins and Filipoviç (2010).

Maximize Frequently Occurring Properties (MaF).

> Properties of the L2 are learned in proportion to their frequency of occurrence (as measured, for example, in the British National Corpus): more frequent exposure of a property to the learner facilitates its learning and reduces learning effort.

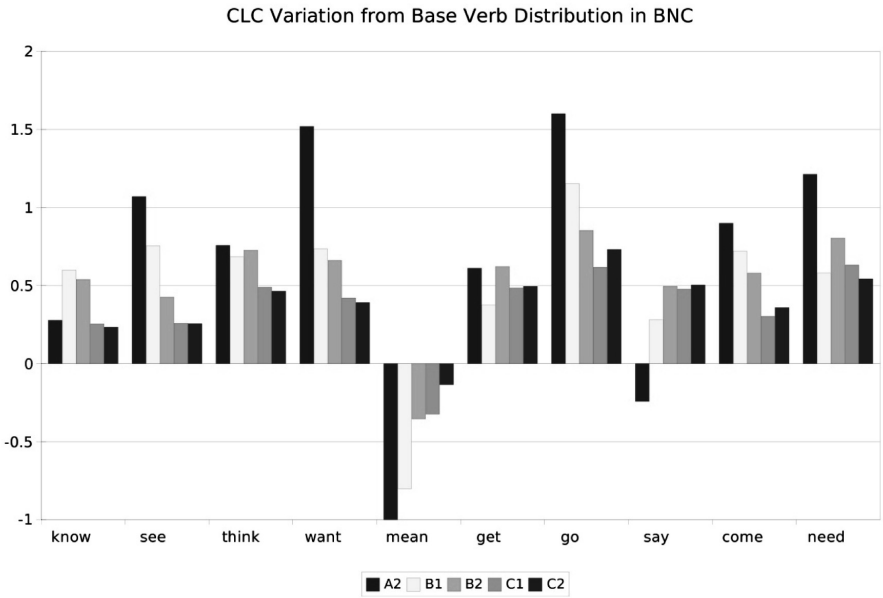
I.e. more frequent properties result in earlier L2 acquisition, more of the relevant properties learned, and fewer errors, in general. Infrequency makes learning more effortful.

In the L1 literature, the importance of frequency of occurrence of a property for its acquisition is attested in the work of Tomasello (2001, 2003) and Diessel (2004) among others. Tomasello (2001, 2003) claims that children's early production consists of specific linguistic expressions that imitate the language they hear around them. It is only later that children creatively combine these expressions to reach adult competence. Tomasello and Brooks (1998), for instance, showed that children younger than 3 years of age use novel verbs only in sentence frames in which they have heard these verbs occurring. In their study, children heard a novel verb (*tam*) for the first time in an intransitive sentence frame, such as *The sock is tamming* (with a meaning similar to *roll* or *spin*). When they were subsequently shown a picture where someone was "tamming" something and were asked *What is X doing?*, most children preferred to re-use the novel verb with the intransitive frame rather than with a transitive one (despite the fact that the question was prompting them to produce a transitive frame). A number of other studies using different constructions corroborate these findings (e.g. Dodson & Tomasello, 1998; Brooks & Tomasello, 1999; Akhtar, 1999). Diessel (2004) also discusses frequency of occurrence in the ambient language as one of the major factors

motivating the order of acquisition of complex sentences in the first language, such as infinitival and participial complement constructions, relative clauses, adverbial and co-ordinate clauses.

In SLA, the work of MacWhinney and colleagues on the Competition Model (Bates & MacWhinney, 1987; MacWhinney, 1987a, 1992, 1997) provides examples of the frequency principle. In the framework of the Competition Model (MacWhinney, 1987b) and more recently of the Unified Model (MacWhinney, 2008), forms (e.g. lexical items) provide cues to functional interpretations for sentence comprehension and conversely, underlying functions provide cues to retrieving forms for sentence production. "Cue availability", i.e. *how often* a particular form occurs as a cue for a certain underlying function, is one major factor in determining "cue validity"[4],

Figure 2: CLC variation from base verb distribution in BNC, including present tense, uninflected forms only from both corpora (reproduced from Hawkins & Buttery, 2009a, p. 164).[5]



4 The other factor is "cue reliability", that is, how reliably a form marks a function.

5 A1 is not included in this figure and the subsequent data tables as the EP findings reported in this paper derive from data ranging from A2 to C2 (cf. Table 1).

which, in turn, determines ease of learning of form/function correspondences. In other words, second language learners are figuring out the form/function mappings of the target language through repeated exposure to these mappings and the frequency of occurrence of these mappings in the input is a decisive factor. Connectionism approaches to SLA also emphasise the role of frequency. In the framework of a connectionist model, second language learning is seen as the extraction of regular patterns from the input. Using a connectionist model, Ellis and Schmidt (1997), for instance, demonstrated frequency effects for the acquisition of L2 morphology, supporting earlier findings in this area by Larsen-Freeman (1976), who argued that frequency of occurrence is a major determiner of the order of acquisition of morphemes in L2.

Turning now to the CLC learner data, one illustration of the MaF principle is the use of the ten most common (frequent) verbs in English (*know, see, think, want, mean, get, go, say, come, need*) by L2 learners in CLC. Hawkins and Buttery (2009a) found that nine out of ten of these verbs are overrepresented in the earlier stages of L2 learning, moving gradually to more native-like L1 English use. Figure 2 shows the ratio of relative frequency of the base form of these ten lexical verbs in the CLC compared with their occurrence in the BNC, taking into account present tense, uninflected forms only in both corpora. Bars above the zero line indicate an over-use in the CLC in comparison to the BNC; bars below the zero line indicate under-use.

Figure 2 shows that overall these common verbs are indeed overrepresented in the CLC relative to the native speaker's usage in the BNC – apart from the verb *mean*. This overrepresentation is a feature of all levels but it declines at the higher level in accordance with the MaF principle above.

Another illustration of MaF in the CLC data comes from verb co-occurrence frames. Williams (2007) analysed verb co-occurrence frames using the Briscoe-Korhonen subcategorisation frame system (cf. Briscoe, 2000; Briscoe & Carroll, 1997; Korhonen, Krymolowski, & Briscoe, 2006; Preiss, Briscoe, & Korhonen, 2007) and recorded the appearance of new frames from A2 to B2 as shown in Tables 3a-3c. There were no new verb co-occurrence frames at C levels suggesting that these basic constructions of English are, more or less, learned by B2. As Hawkins and Buttery (2009a) remark, C2 levels require a different and more subtle kind of analysis in order to capture progress, and projects are planned to explore this issue.

Table 3a: New verb co-occurrence frames at A2 level (Williams, 2007)

Frame	Example
• NP-V	*He went*
• NP-V (reciprocal Subj)	*They met*
• NP-V-PP	*They apologized* [*to him*]
• NP-V-NP	*He loved her*
• NP-V-Part-NP	*She looked up* [*the number*]
• NP-V-NP-Part	*She looked* [*the number*] *up*
• NP-V-NP-PP	*She added* [*the flowers*] [*to the bouquet*]
• NP-V-NP-PP (P = *for*)	*She bought* [*a book*] [*for him*]
• NP-V-V(+ *ing*)	*His hair needs combing*
• NP-V-VPinfinitival (Subj Control)	*I wanted to play*
• NP-V-S	*They thought* [*that he was always late*]

Table 3b: New verb co-occurrence frames at B1 level (Williams, 2007)

Frame	Example
• NP-V-NP-NP	*She asked him* [*his name*]
• NP-V-Part	*She gave up*
• NP-V-VPinfin (WH-move)	He explained [*how to do it*]
• NP-V-NP-V(+ *ing*) (Obj Control)	*I caught him stealing*
• NP-V-NP-PP (P = *to*) (Subtype: Dative Movement)	*He gave* [*a big kiss*] [*to his mother*]
• NP-V-NP-(*to be*)-NP (Subj to Obj Raising)	*I found him (to be) a good doctor*
• NP-V-NP-Vpastparti (V = passive) (Obj Control)	*He wanted* [*the children*] *found*
• NP-V-P-V*ing*-NP (V = + *ing*) (Subj Control)	*They failed in attempting the climb*
• NP-V-Part-NP-PP	*I separated out* [*the three boys*] [*from the crowd*]
• NP-V-NP-Part-PP	*I separated* [*the three boys*] *out* [*from the crowd*]
• NP-V-S (Wh-move)	*He asked* [*how she did it*]
• NP-V-PP-S	*They admitted* [*to the authorities*] [*that they had entered illegally*]
• NP-V-S (*whether* = Wh-move)	*He asked* [*whether he should come*]
• NP-V-P-S (*whether* = Wh-move)	*He thought about* [*whether he wanted to go*]

Table 3c: New verb co-occurrence frames at B2 level (Williams, 2007)

Frame	Example
• NP-V-NP-AdjP (Obj Control)	*He painted* [*the car*] *red*
• NP-V-NP-*as*-NP (Obj Control)	*I sent him as* [*a messenger*]
• NP-V-NP-S	*He told* [*the audience*] [*that he was leaving*]
• NP-V-P-NP-V(+*ing*) (Obj Control)	*They worried about him drinking*
• NP-V-VPinfin (Wh-move)(Subj Control)	*He thought about* [*what to do*]
• NP-V-S (Wh-move)	*He asked* [*what he should do*]
• NP-V-Part-VPinfin (Subj Control)	*He set out to win*

Critically, Williams found that the progression from A2 to B2 correlates with the frequency of these frames in native speaker corpora – the most frequent frames appear at A2 moving progressively to less frequent frames at B1 and B2. Table 4 provides the average token frequencies in BNC for the subcategorisation frames identified as occurring for the first time at A2-B2 in Tables 3a-c.

Table 4: Average token frequencies in native English corpora (BNC) for the new verb co-occurrence frames appearing from A2-B2 in the CLC (cf. Tables 3a-c).

CEFR Level	A2	B1	B2
Average token frequency	1,041,634	38,174	27,615

Maximise Structurally and Semantically Simple Properties (MaS)

> Properties of the L2 are learned in proportion to their structural and semantic simplicity: simplicity means that there are fewer properties to be learned and less learning effort is required. Simpler properties result in earlier L2 acquisition, more of the relevant properties learned, and fewer errors. Complexity makes learning more effortful, in general, since there are more properties to be learned.

In addition:

> Properties of the L2 are used in proportion to their structural and semantic simplicity: simplicity means that there are fewer properties to be processed in on-line language use and less processing effort is required.

That is, simplicity and complexity affect both learning and processing. This principle then predicts that simpler constructions will be acquired earlier than more complex ones. But what constitutes a simple or a complex structure? There is not as yet a satisfactory account of complexity in SLA research (see Rimmer, 2006, for a review of complexity in SLA) and this is why we turn to L1 and typology research. Hawkins (2004) defines complexity as:

> Complexity increases with the number of linguistic forms and the number of conventionally associated (syntactic and semantic) properties that are assigned to them when constructing syntactic and semantic representations for sentences. That is, it increases with more forms, and with more conventionally associated properties. It also increases with larger formal domains for the assignment of these properties. (p. 9; see also Hawkins, 2009)

The different types of relative clauses as listed in the Keenan-Comrie Accessibility Hierarchy Hypothesis (AH; 1977) provide a good example of increasing complexity across structures. Table 5 below lists types of relative clauses in order of complexity following the Keenan-Comrie AH and Hawkins (2004). Hawkins (1994, pp. 37-42, 2004, pp. 177-8, 2009) has argued that the relative clauses down the AH involve increasing complexity of the processing domains for the different relativizable positions. According to Hawkins, the number of dominating and co-occurring syntactic nodes required for these relativizable positions correlates strongly with their AH ranking – the lower the ranking the higher the number of syntactic nodes required and the greater the complexity of the relative clause. (For a detailed account of this argument, the reader is referred to Hawkins, 1994, 2004, 2009).

Table 5: Relative clause types in order of complexity based on the Keenan-Comrie Accessibility Hierarchy (1977) and Hawkins (2004)

Relative Clause Types *(in order of complexity based on the Keenan-Comrie Accessibility Hierarchy, 1977, and Hawkins, 2004)*	
Subject Relatives	*The student **who/that** wrote the paper*
Direct Object Relatives	*The student **who(m)/that** I taught*
Indirect/Oblique Object Relatives	*The student **to whom** I gave the book* *The student **who/that** I gave the book **to***
Genitive Relatives (within a Subject)	*The student **whose** supervisor retired*
Genitive Relatives (within a Direct Object)	*The student **whose** supervisor I know*
Genitive Relatives (within an Indirect Object)	*The student **to whose** memory I devoted the book* *The student **whose** memory I devoted the book **to***

According to the MaS principle above then, simpler relative clause types, such as subject relatives, will be learned earlier than more complex ones, such as relatives in genitive positions (see Table 5). And this is what is actually attested in the CLC data (see Table 6 below). Hawkins and Buttery (2009b) found a clear progression from A2 to C2 in the appearance of new relative clause types. This progression correlates with the increasing complexity of the relative clauses involved (subject > object > indirect/oblique > genitive; cf. Keenan & Comrie, 1977; Hawkins, 1994). For example, fairly complex relative clauses from the lower positions of the Keenan-Comrie AH (1977), such as relatives in indirect/oblique object positions (*the student to whom I gave the book*), are rare

before B2, whereas relatives in genitive positions (*the student whose supervisor retired*) do not appear at all before C1.[6]

Table 6: Usage of different types of relative clauses as percentage of total within each CEFR level

	A2	B1	B2	C1	C2
Subject RCs	67.7%	61.1%	71.1%	70.3%	74.4%
Direct Object RCs	30.7%	37.3%	26.1%	25.0%	20.9%
Indirect Object RCs	1.6%	1.6%	2.8%	4.6%	4.3%
Genitive RCs	0.0%	0.0%	0.0%	0.1%	0.2%

These findings are in accordance with a substantial body of SLA research on the acquisition of relative clauses by second language learners from various L1 backgrounds which show that the order of acquisition appears to follow the ranking of the AH (e.g. Doughty, 1991; Eckman, Bell, & Nelson, 1988; Gass 1979, 1980; Gass & Ard, 1984; O'Grady, Lee, & Choo, 2003). Gass (1979, 1980), for example, found that the production of different types of relative clauses by learners of English with a wide range of L1s could be predicted based on their rank order in the AH – that is, the percentage correct of subject relatives was higher than the percentage correct of direct object relatives, and so on. These findings were based on empirical data from a variety of experimental tasks, including free compositions, sentence combining and grammaticality judgements. (But see also the more recent studies of Jeon & Kim, 2007; Ozeki & Shirai, 2007, for some exceptions.)

Lexical semantic properties appear to follow a similar acquisitional path in L2: simpler semantic senses are learned earlier than more complex, figurative senses (Hawkins & Filipoviç, 2010). Table 7 illustrates this principle by looking at the acquisition of *break*.[7]

---

6 In all levels, the learners were responding to open-ended, essay type questions (with minimal prompts) that allowed considerable freedom of expression in terms of content and form. It is thus unlikely that the distribution in Table 6 is due to what the writing tasks in the different CEFR levels allowed the learners actually to use rather than progressive learning across the levels. However, this remains an empirical question that will be further investigated as additional learner data from non-exam settings are collected and analysed within EP (see the Conclusions section).

7 These findings deriving from learner corpora appear to be in accordance with earlier work on prototypicality and language transfer by Kellerman (1977, 1978, 1979).

Table 7: Occurrence of break in the CLC

CEFR level	Example of occurrence	Type of use
A2	break	basic physical sense
B1	break the routine	additional sense of INTERRUPT
B2	break an argument / promise	additional sense of NOT OBEY
C1	break the bank	Idiomatic
C2	break the wall that surrounds him	original figurative

Maximise Positive Transfer (MaPT)

> Properties of the L1 which are also present in the L2 are learned more easily and with less learning effort, and are readily transferred, on account of pre-existing knowledge in L1.

Similar or identical L1/L2 properties will result in earlier L2 acquisition, more of the relevant properties learned, and fewer errors, in general, unless these shared properties are impacted by other factors, such as high complexity (cf. MaS) or low frequency of occurrence (cf. MaF). Dissimilar L1/L2 properties will be harder to learn by virtue of the additional learning that is required, in general, and this learning may be more or less effortful depending on other factors (cf. Odlin, 2005, for a summary of relevant research literature in SLA). With respect to errors, dissimilar L1/L2 properties will result either in more errors or in structural avoidance (and hence possibly in fewer errors, e.g. Schachter, 1974). The more obligatory or unavoidable the lexical/grammatical property in question, the more we will see errors rather than avoidance.

A test for the above principle is the use of definite and indefinite articles in L2 English by learners whose first language has an article system and learners whose first language does not use articles. MaPT predicts that the acquisition of

---

In a series of studies, Kellerman investigated learners' intuitions about which of the different meanings of the Dutch verb *breken* can be 'transferred' in English, i.e. translated into its English cognate *break*. He found that the dimension of 'prototypicality' largely determined Dutch learners' judgements in that more core senses of *breken* (e.g. "He broke his leg", "The cup broke") presented higher percentages of transferability than less core, figurative senses (e.g. "The underground resistance was broken", "A game would break up the afternoon a bit"). This order of 'transferability' closely resembles the order of acquisition of the senses of *break* in English as L2 identified in the present study.

the article system of English will be easier for the former group of learners. And this is what is attested in the CLC data (Hawkins & Buttery, 2009a). Table 8 displays error rates for *the* (definite) and *a* (indefinite) articles at CEFR levels A2-C2 by French, German and Spanish learners of English, whose first language has a similar article system to that of English. The numbers are percentages of errors compared to the total number of correct uses. As Table 8 shows, error rates are generally low for these learners without significant differences between the CEFR levels (Hawkins & Buttery, 2009a).

Table 8: Missing Determiner Error Rates for L1s with Articles (Hawkins & Buttery, 2009a, p.168)

Missing 'the'					
	A2	**B1**	**B2**	**C1**	**C2**
French	4.76	4.67	5.01	3.11	2.13
German	0.00	2.56	4.11	3.11	1.60
Spanish	3.37	3.62	4.76	3.22	2.21
Missing 'a'					
	A2	**B1**	**B2**	**C1**	**C2**
French	6.60	4.79	6.56	4.76	3.41
German	0.89	2.90	3.83	3.62	2.02
Spanish	4.52	4.28	7.91	5.16	3.58

Now compare the data in Table 8 above with those displayed in Table 9 below which shows error rates for the same articles by Turkish, Japanese, Korean, Russian and Chinese learners of English. These languages do not have an article system. Error rates are significantly higher across all CEFR levels than error rates at the equivalent levels by learners with first languages which have articles. However, unlike learners whose first language has articles, learners whose first language has no articles show, in general, a linear improvement, i.e. a decline in error rates as they progress through the CEFR levels.[8] (A more detailed study

8 The only exception are the Chinese learners who present an inverted U-shaped pattern of error rates, particularly in the use of a, with significant improvement only at C2.

(Alexopoulou, 2010) on definiteness and indefiniteness and the various ways in which articles (including zero article) are used across the different CEFR levels is currently underway in EP).

Table 9: Missing Determiner Error Rates for L1s without Articles (Hawkins & Buttery, 2009a, p. 169)

Missing 'the'					
	A2	**B1**	**B2**	**C1**	**C2**
Turkish	22.06	20.75	21.32	14.44	7.56
Japanese	27.66	25.91	18.72	13.80	9.32
Korean	22.58	23.83	18.13	17.48	10.38
Russian	14.63	22.73	18.45	14.62	9.57
Chinese	12.41	9.15	9.62	12.91	4.78
Missing 'a'					
	A2	**B1**	**B2**	**C1**	**C2**
Turkish	24.29	27.63	32.48	23.89	11.86
Japanese	35.09	34.80	24.26	27.41	15.56
Korean	35.29	42.33	30.65	32.56	22.23
Russian	21.71	30.17	26.37	20.82	12.69
Chinese	4.09	9.20	20.69	26.78	9.79

These findings are in line with an abundance of earlier studies that report L1 transfer effects in SLA – not only for articles (e.g. Dušková, 1983) but for all aspects of language learning (for a comprehensive review of studies on L1 transfer and cross-linguistic influence see e.g. Odlin, 2005; Gass & Selinker, 1992, 2008). There are also numerous transfer theories and models, ranging from generativist ones (e.g. Full Transfer/Full Access Model, Schwartz & Sprouse, 1996) to processing approaches (e.g. Developmentally Moderated Transfer Hypothesis, Pienemann, Di Biase, Kawaguchi, & Håkansson, 2005; L1 and L2 cue competition, MacWhinney, 1992, 2008) and connectionist approaches (Ellis, 2008). The main questions investigated by these theories include the extent of availability of L1 features during the initial and later stages of SLA, and the interaction of or competition between L1 and L2 features. It is beyond the scope of this chapter and EP to provide a full review of existing transfer the-

ories or test their predictions, as explained in the Introductory section. Odlin (2005) remarks that "[t]he highly diverse evidence for transfer has impeded attempts to develop truly comprehensive theories of cross-linguistic influence. In the more credible attempts at theory-building, researchers have focused on what is admittedly only part of an overall model" (p. 437). The EP goal is to account for L1 transfer as one of the factors within a multi-factor model of SLA using a complex adaptive system and computer simulation (see the concluding section for more details).

With respect to the identification of criterial features, the aim of EP research is two-fold. First, it aims at identifying criterial features that are not L1-specific as one of the main aims of the CEFR is to compare L2 learners across their L1s. Examples of such features are provided under the MaF and MaS principles above. However, as the cross-linguistic data in the last section indicate, cross-linguistic variation does exist among L2 learners across the CEFR levels, and in some cases, e.g. the omission of articles, this variation persists up to C2 level. EP research, therefore, also aims at investigating to what extent criterial features and thus linguistic profiles per CEFR level may differ depending on the L1 of the learner. Is it the case that different groups of learners make use of different criterial features and as a result have a substantially different linguistic profile per CEFR level according to their L1? Is the description of *one single* general linguistic profile per CEFR level viable or should extensive reference be made to subgroups of L1s? These are issues currently under investigation in EP in line with SLATE's objectives (cf. research question 2 in Introductory chapter, this volume).

Moreover, the gradual transition in recent years from traditional to more diverse learning settings around the globe provides an additional impetus for the investigation of cross-linguistic differences across the CEFR levels. In a traditional learning setting (e.g. classroom), the teacher would typically address the needs of a homogeneous group; in a modern learning setting (e.g. an e-learning environment where learners from around the world meet virtually in chat rooms), teachers would cater to learners from a wide range of linguistic backgrounds with diverse learning needs. This gives rise to a new need for learners: personalised learning. Linguistic research on the effect of the L1 can inform teaching and address requirements for personalised learning, of increased importance in settings of globalised e-learning environments (Alexopoulou, Yannakoudakis, & Briscoe, 2010).

In summary, this section illustrated three principles that appear to drive SLA development in the CLC data. This is work in progress and current EP research is formulating further principles to account as comprehensively as possible for the emerging learning patterns identified in CLC (for more details see

Hawkins & Filipović, 2010). It needs to stressed, however, that we do not claim that any of the above principles alone (or any single learning principle for that matter) can fully account for SLA development and performance independently of one another and/or other factors. In fact, further EP research now focuses on identifying the interactions between these principles and their predictive power within a multi-factor SLA model (see *Conclusions and the way forward* below for specific proposals of how this investigation will take shape).

6. 'Criterial features' of English across the CEFR levels

The last two sections outlined some language features evident across the CEFR levels as illustrations for the L2 principles and patterns identified thus far. For a more comprehensive inventory of the criterial features identified so far (on the basis of the principles described above), the reader is referred to Salamoura and Saville (2009). A more complete list is currently developed by Hawkins and Filipović (2010). The list of criterial features has also been informed by earlier research on the properties of learner English at different learning stages, namely the T-series (*Breakthrough*, Trim, 2001; *Waystage*, *Threshold* and *Vantage*, van Ek & Trim, 1998a, 1998b, 2001). Researchers within the EP team are currently revisiting these publications in search of features that are novel at each level and that could thus qualify for the status of criterial features. Some preliminary results from this project are again provided in Salamoura and Saville (2009).

It should be stressed that these preliminary findings are a 'snapshot' of EP research as it stands at the time this chapter goes to press. It is expected that these findings will be refined, revised and complemented as more data become available and as more research is carried out.

7. Conclusion and the way forward

In summary, in this chapter we discussed EP's approach to profiling the CEFR levels – a search for criterial features of learner performance that distinguish the CEFR levels, informed by SLA theory and empirically derived from an extensive L2 English corpus, the CLC, using psycholinguistic and computational principles and metrics. The findings discussed in this chapter are mostly preliminary but they already reveal a promising picture of a learner's profile. A number of criterial features have emerged across the CEFR levels which show, from A2 to C2, an increasing progression in terms of frequency, syntactic and semantic complexity, and at the same time, a decreasing tendency for errors, non-native like usage and L1 influence. Critically, this progression can be systematically quantified and

measured against a large corpus of learner data. The emerging performance patterns per CEFR level are potentially highly informative for our understanding of the development of SLA, as they can inform us about the order of acquisition of linguistic features and elucidate the role and interaction of factors such as frequency, complexity and L1 transfer. It thus appears that the EP approach, which reflects closely SLATE's goals and objectives, is fulfilling the original EP aim, which is to produce an exemplification of the CEFR following SLA theory and at the same time shed light on questions and issues raised by SLA and psycholinguistic theory based on empirical (learner) data.

Although quite informative, these initial findings raise a number of further questions which EP researchers are currently addressing. These include:

- How do the different SLA patterns and principles that emerge from the study of CLC interrelate? This is a particularly important question as some of the emerging principles in CLC may be in competition with each other. Which competing principle takes precedence over the other? Hawkins and Filipoviç (2010) argue that the principles of frequency, complexity and L1 transfer can be incorporated within a multi-factor model of SLA and used to define possible versus impossible, and likely versus unlikely, IL stages. They propose to investigate the relative strength and interaction between principles by setting up a computer simulation that defines these possible/impossible and likely/unlikely IL stages in the manner of a complex adaptive system (Gell-Mann, 1992), and in the manner of Kirby's (1999) computer simulation of the emergence of possible/impossible and likely/unlikely word order variants, using the processing principles of Hawkins (1994).
- How do the different kinds of criterial features (lexical semantic, morpho-syntactic, syntactic, discourse, notional, functional, etc.) interrelate? In particular, which linguistic features realise which language functions across the CEFR levels? Answering these questions is fundamental in bringing together the different strands of EP research.
- To what extent does the criteriality of features vary depending on the L1 of the learner?
- Which criterial features can be used as diagnostics at the individual learner level?
- What is the effect of task type on learner production and criterial features? (Parodi, 2008)
- How does the type of context in which some linguistics properties (e.g., spatial verbs) occur help explain the emerging patterns in CLC (Hendriks, 2008)?

The immediate future of the EP will involve extending the current analyses to broader samples from the CLC and collecting other kinds of non-exam written data (e.g. from classroom settings) from learners of English worldwide. A major data collection exercise is currently being undertaken worldwide to this effect. Another major challenge being addressed is how to include *spoken language* in the analysis (McCarthy & Saville, 2009) in order to be able to describe a learner's linguistic profile at each CEFR level for speaking too. Such a profile will complement the current linguistic profile being investigated for writing (as defined in terms of lexico-semantic, morphosyntactic and syntactic features listed on pp. 6–7 and exemplified through the SLA principles identified thus far in the CLC data: MaF, MaS and MaPT). This would bring the CEFR profiling a step closer to what SLATE envisages as a complete linguistic description of the CEFR for all four skills (see research question 1 in Introductory chapter, this volume). Finally, another future aim is to collect data that will make it possible to foster a closer relationship between the EP outcomes and teachers/learners of English in their different contexts world-wide (Alexopoulou, 2008).

The method described in this chapter for profiling the CEFR ('criterial features', SLA theory- and corpus-informed empirical approach) was illustrated for the English language. However, once fully developed and established, it has the potential for application to any second language (Hendriks, 2008). Such a prospect would without doubt facilitate the comparison of CEFR profiles across different L2s as recommended by the SLATE network (see research question 2 in Introductory chapter, this volume).

It is envisaged that the description of English across the CEFR levels in terms of criterial features will result in a valuable data source for researchers and a useful tool for practitioners in the fields of English language learning, teaching and assessment. Moreover, as an outcome of the EP, it is hoped that the CEFR itself can be operationalised more effectively for English and that it will become a more useful tool for its intended purposes. The search for criterial features will lead to better *linguistic descriptions*, and this in turn will lead to better *functional descriptors*, thus addressing a current weakness (see Milanovic, 2009). Already the focus on empirical research at the bottom and top ends of the scale (A1, and C1/2) is providing more precise information about the nature of proficiency in English at these levels.

Acknowledgements

We are grateful to Prof. John Hawkins, Dr. Luna Filipović, Dr. Teresa Parodi, Dr. Henriette Hendriks and Prof. Roger Hawkey for their input in an earlier version of this chapter.

References

Akhtar, N. (1999). Acquiring basic word order: Evidence for data-driven learning of syntactic structure. *Journal of Child Language, 26*, 261–278.

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

Alexopoulou, T. (2010). *"She's great teacher": Article omission in predicative sentences or the non-linearity of L1 effects on English L2 predicative nominals*. Manuscript in preparation.

Alexopoulou, T., Yannakoudakis, H. & Briscoe, E. J. (2010). *L1 effects and personalised learning in globalised learning settings*. Manuscript in preparation.

Bates, E., & MacWhinney, B. (1987). Competition, variation, and language learning. In B. MacWhinney (Ed.), *Mechanisms of language acquisition* (pp. 157–194). Hillsdale, NJ: Lawrence Erlbaum Associates.

Bialystok, E. (2001). *Bilingualism in development: Language, literacy and cognition.* Cambridge: Cambridge University Press.

Briscoe, E. J. (2000). *Dictionary and system subcategorisation code mappings*. MS, Computer Laboratoty, University of Cambridge.

Briscoe, E. J., & Carroll, J. (1997). Automatic extraction of subcategorisation from corpora. In *Proceedings of the 5th ACL Conference on Applied Natural Language Processing* (pp. 356–363). Washington DC. Retrieved from http://delivery.acm.org/10.1145/980000/974609/p356-briscoe.pdf?key1=974609&key2=8461832721&coll=GUIDE&dl=GUIDE&CFID=88116625&CFTOKEN=35472625

Briscoe, E. J., Carroll, J., & Watson, R. (2006). The second release of the RASP system. In *Proceedings of the COLING/ACL 2006 Interactive Presentation Sessions* (pp. 77–80). Sydney, Australia. Retrieved from http://acl.ldc.upenn.edu/P/P06/P06-4020.pdf

Brooks, P., & Tomasello, M. (1999). Young children learn to produce passives with nonce verbs. *Developmental Psychology, 35*, 29–44.

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

Corder, S. P. (1967). The significance of learners' errors. *International Review of Applied Linguistics, 5*, 161–170.

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

Diessel, H. (2004). *The acquisition of complex sentences.* Cambridge: Cambridge University Press.

Dodson, K., & Tomasello, M. (1998). Acquiring the transitive construction in English: The role of animacy and pronouns. *Journal of Child Language, 25*, 555–574.

Doughty, C. (1991). Second language instruction does make a difference. *Studies in Second Language Acquisition, 13*, 431–469.

Doughty, C. J., & Long, M. H. (Eds.). (2005). *The handbook of second language acquisition*. Malden, MA: Blackwell.

Dušková, L. (1983). On sources of errors in foreign language learning. In B. Robinett & J. Schachter (Eds.), *Second language learning: Contrastive analysis, error analysis, and related aspects* (pp. 215–233). Ann Arbor: University of Michigan Press.

Eckman, F., Bell, L., & Nelson, D. (1988). On the generalization of relative clause instruction in the acquisition of English as a second language. *Applied Linguistics, 9*, 1–13.

Ellis, N. C. (2008). Usage-based and form-focused language acquisition: The associative learning of constructions, learned attention and the limited L2 endstate. In P. Robinson & N. C. Ellis (Eds.), *Handbook of Cognitive Linguistics and Second Language Acquisition* (pp. 372–405). New York: Routledge.

Ellis, N. C., & Schmidt, R. (1997). Morphology and longer distance dependencies: Laboratory research illuminating the A in SLA. *Studies in Second Language Acquisition, 19*, 145–171.

Gass, S. M. (1979). Language transfer and universal grammatical relations. *Language Learning, 29*, 327–344.

Gass, S. M. (1980). An investigation of syntactic transfer in adult second language learners. In R. Scarcella & S. Krashen (Eds.), *Research in second language acquisition* (pp. 132–141). Rowley, MA: Newbury House.

Gass, S. M., & Ard, J. (1984). Second language acquisition and the ontology of language universals. In W. E. Rutherford (Ed.), *Language universals and second language acquisition* (pp. 33–67). Amsterdam: John Benjamins.

Gass, S. M., & Selinker, L. (Eds.). (1992). *Language transfer in language learning.* Amsterdam: John Benjamins.

Gass, S. M., & Selinker, L. (2008). *Second language acquisition: An introductory course* (3rd ed.). New York / London: Routledge.

Gell-Mann, M. (1992). Complexity and complex adaptive systems. In J.A. Hawkins & M. Gell-Mann (Eds.), *The evolution of human languages* (pp. 3–18). Redwood City, CA: Addison Wesley.

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

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

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

Hawkins, J. A. (2009). An efficiency theory of complexity and related phenomena. In D. Gill, G. Sampson & P. Trudgill (Eds.), *Complexity as an evolving variable* (pp. 252–268). Oxford: Oxford University Press.

Hawkins, J. A., & Buttery, P. (2009a). 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* (pp. 158–175). Cambridge: Cambridge University Press.

Hawkins, J.A., & Buttery, P. (2009b). *Criterial features in learner corpora: Theory and illustrations.* Manuscript submitted for publication.

Hawkins, J. A., & Filipović, L. (2010). *Criterial features in the learning of English: Specifying the reference levels of the Common European Framework.* Manuscript in preparation.

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

Hulstijn, J. H. (2006, June). *Defining and measuring the construct of second language proficiency*. Plenary paper presented at the joint conference of AAAL and ACLA/CAAL, Montreal.

Jeon, K. S., & Kim, H-Y. (2007). Development of relativization in Korean as a foreign language: Noun phrase accessibility hierarch in head-internal and head-external relative clauses. *Studies in Second Language Acquisition, 29*, 253–276.

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

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

Jones, N. (2002). Relating the ALTE framework to the Common European Framework of Reference. In J. C. Alderson (Ed.), *Common European Framework of Reference for Languages: Learning, teaching, assessment – Case studies* (pp. 167–83). Strasbourg: Council of Europe. Retrieved from www.coe.int/T/DG4/Portfolio /documents/case_studies_CEF.doc

Keenan, E. L., & Comrie, B. (1977). Noun phrase accessibility and Universal Grammar. *Linguistic Inquiry, 8*, 63–99.

Kellerman, E. (1977). Towards a characterization of the strategy of transfer in second language learning. *Interlanguage Studies Bulletin, 2*, 58–145.

Kellerman, E. (1978). Giving learners a break: Native language intuitions as a source of predictions about transferability. *Working Papers on Bilingualism, 15*, 59–92.

Kellerman, E. (1979). Transfer and non-transfer: Where are we now? *Studies in Second Language Acquisition, 2*, 37–57.

Kessler J-U. (2007). Assessing EFL-Development online: A feasibility study of Rapid Profile. In F. Mansouri (Ed.), *Second language acquisition research: Theory-construction and testing* (pp. 119–144). Newcastle: Cambridge Scholars.

Kirby, S. (1999). *Function, selection and innateness: The emergence of language universals.* Oxford: Oxford University Press.

Korhonen, A., Krymolowski, Y., & Briscoe, E. J. (2006). A large subcategorization lexicon for natural language processing applications. In *Proceedings of the 5th International Conference on Language Resources and Evaluation* (LREC06), Genova, Italy. Retrieved from http://www.cl.cam.ac.uk/~alk23/lrec06-lexicon.pdf

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

Larsen-Freeman, D. (1976). An explanation for the morpheme acquisition order of second language learners. *Language Learning, 26*, 125–134.

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*, 645–55.

MacWhinney, B. (1987a). Applying the competition model to bilingualism. *Applied Psycholinguistics, 8*, 315–327.

MacWhinney, B. (1987b). The Competition Model. In B. MacWhinney (Ed.), *Mechanisms of language acquisition* (pp. 249–308). Hillsdale, NJ: Lawrence Erlbaum Associates.

MacWhinney, B. (1992). Competition and transfer in second language learning. In R. J. Harris (Ed.), *Cognitive processing in bilinguals* (pp. 371–390). Amsterdam: Elsevier Science.

MacWhinney, B. (1997). Second language acquisition and the Competition Model. In A. de Groot & J. Kroll (Eds.), *Tutorials in bilingualism* (pp. 113–142). Mahwah, NJ: Lawrence Erlbaum Associates.

MacWhinney, B. (2008). A Unified Model. In N. Ellis & P. Robinson (Eds.), *Handbook of cognitive linguistics and second language acquisition* (pp. 341–371). New York: Routledge.

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

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

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

O'Grady, W., Lee, M., & Choo, M. (2003). A subject-object asymmetry in the acquisition of relative clauses in Korean as a second language. *Studies in Second Language Acquisition, 25*, 433–448.

Odlin, T. (2005). Cross-linguistic influence. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 436–489). Malden, MA: Blackwell.

Ozeki, H., & Shirai, Y. (2007). The consequences of variation in the acquisition of relative clauses: An analysis of longitudinal production data from five Japanese children. In Y. Matsumoto, D. Oshima, O. Robinson, & P. Sells (Eds.), *Diversity in language: Perspectives and implications.* Stanford, CA: CSLI Publications.

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

Pienemann, M. (1992). COALA – A computational system for interlanguage analysis. *Second language Research, 8*, 59–92.

Pienemann, M., Di Biase, B., Kawaguchi, S., & Håkansson, G. (2005). Processing constraints on L1 transfer. In J. F. Kroll & A. M. B. DeGroot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 128–153). New York: Oxford University Press.

Pienemann, M., & Keßler J-U. (2007). Measuring bilingualism. In P. Auer & L. Wei, (Eds.), *Handbook of multilingualism and multilingual communication* (pp. 245–275). Berlin: Mouton de Gruyter.

Preiss, J., Briscoe, E. J., & Korhonen, A. (2007). A system for large-scale acquisition of verbal, nominal and adjectival subcategorization frames from corpora. In

Proceedings of the 45th Annual Meeting of the Association for Computational Linguistics (pp. 912–919), June 2007, Prague, Czech Republic. Retrieved from http://www.cl.cam.ac.uk/~alk23/acl-07.pdf

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

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

Schachter, J. (1974). An error in error analysis. *Language Learning, 24*, 205–214.

Schwartz, B., & Sprouse, R. (1996). L2 cognitive states and the Full Transfer/Full Access model. *Second Language Research, 12*, 40–72.

Skehan, P. (1998). *A cognitive approach to language learning.* Oxford: Oxford University Press.

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

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

Tomasello, M. (2001). The item-based nature of children's early syntactic development. In M. Tomasello & E. Bates (Eds.), Language development: The essential readings (pp. 169–202). Oxford: Blackwell Publishing.

Tomasello, M. (2003). Constructing a language: A usage-based theory of language acquisition. Cambridge, MA: Harvard University Press.

Tomasello, M., & Brooks, P. (1998). Young children's earliest transitive and intransitive constructions. *Cognitive linguistics, 9*, 379–395.

Trim, J. L. M. (2001). *Breakthrough.* Unpublished manuscript. Retrieved from www.englishprofile.org

Van Ek, J., & Trim, J. L. M. (1998a). *Threshold 1990.* Cambridge: Cambridge University Press. (Original work published 1991).

Van Ek, J., & Trim, J. L. M. (1998b). *Waystage 1990.* Cambridge: Cambridge University Press. (Original work published 1991).

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

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

Young, R. (1991). *Variation in interlanguage morphology.* New York: Peter Lang.