

## ACCURACIES AND INACCURACIES IN EFL LEARNERS' WRITTEN VOCABULARY USE<sup>1</sup>

LEANNE BARTLEY\*  
MIGUEL-ÁNGEL BENÍTEZ-CASTRO  
*University of Granada (Spain)*

**ABSTRACT.** *This paper draws on a longitudinal corpus-based study into the lexical competence of first-year Spanish university students of English, based on a set of written compositions collected over two semesters within the same academic year (2003-2004). The methodology employed for this paper takes account of the correct and incorrect tokens produced by students, as well as of students' lexical variety (i.e. Lexical Frequency Profile) as compared to a set of frequency lists retrieved from the British National Corpus (BNC). The analysis for this paper was carried out manually. The results reveal an overall minimal progression over an academic year, this pointing to the potential need to focus on a longer time-span in the evolution of students' interlanguage.*

**KEY WORDS.** *Learner language, Lexis, LFP, Error analysis.*

**RESUMEN.** *Este artículo se basa en un estudio de corpus longitudinal sobre la competencia léxica de alumnos universitarios españoles de primer curso de Filología Inglesa. El estudio se centra en el análisis de un muestreo recogido a lo largo de dos semestres dentro del mismo curso académico (2003-2004). Se emplea una metodología que identifica, por una parte, el grado de variedad en el vocabulario utilizado, (i.e. Lexical Frequency Profile), además de su nivel de precisión, en comparación con un grupo de listas de frecuencia recogidas del British National Corpus (BNC). El análisis se llevó a cabo manualmente. Los resultados revelan una progresión mínima a lo largo de un año académico, lo que parece indicar la necesidad potencial de centrarse en un periodo de tiempo de mayor longitud en la evolución de la interlengua de los alumnos.*

**PALABRAS CLAVE.** *Lenguaje de los aprendices, Léxico, LFP, análisis de errores.*

### 1. INTRODUCTION

Traditional research into second language acquisition has focused on L2 learners' deviant use of the language. Error analysis, epitomized in the works by Corder (1967),

James (1998) and Ellis and Barkhuizen (2005), on its own reveals only part of the picture. Learning a language is not solely a matter of fixing incorrect uses of linguistic units, but it also involves acquiring its vocabulary in both its passive (comprehension) and active (production) facets. As argued by Laufer (1986: 69), “[n]o language acquisition, whether first, second, or foreign; child, or adult, can take place without the acquisition of lexis”. Based on this idea, some research has been conducted in order to establish the range, sophistication and depth of the lexical items foreign language learners employ in their written output (e.g. Laufer 1994; Laufer and Nation 1995; Nation 2001).

This paper reviews the variety and accuracy of EFL students’ productive use of written vocabulary, based on a sample of texts (12,500 words) written by thirty Spanish first-year undergraduates at two different stages of an academic year. This article was intended with two goals in mind: on the one hand, to examine the lexical range in learner writing, and on the other, to determine the accuracy of vocabulary use in the two samples considered. The latter entailed the classification of tokens into the six error categories comprising the *EARS*<sup>2</sup> error tagging system (Díaz-Negrillo 2009), namely Spelling, Lexis, Punctuation, Clause grammar, Phrase grammar and Word grammar. As such, the approach to the analysis of lexical errors in this paper comprises cases of formal or semantic infelicities as well as errors derived from the incorrect combination of items along the syntagmatic axis (clause and phrase). Lastly, the analysis focused on one aspect of certain error tags, namely the dichotomy between external and internal errors (see 3.2 for details).

Ideally, the results from this study should cast light on the evolution in the complexity of the subjects’ lexical range over one academic year. However, as suggested by Engber (1995: 50), students may well be at a stage where they have not yet reached optimum lexical competence and, thus, their improvement could be either very slight or non-existing. This paper provides figures of the actual lexical improvement measured as described above.

## 2. THEORETICAL BACKGROUND

### 2.1. *Lexical measures*

Non-native writing, often regarded as a deviation from native standards, is at the centre of a wide range of scholarly research. Substantial differences have emerged from analyses comparing L1 and L2 writing, particularly with regard to lexical variety and syntactic choices (Silva 1993, cited in Ferris 2001: 299), to the extent that L2 learners’ difficulty with writing in the foreign language is often reported to result from their insufficient lexical competence (Laufer 1986; Granger 1998; Hinkel 2002). In this respect, a number of references point to the correlation between lexical richness and the quality of the students’ written output (Engber 1995; Laufer and Nation 1995; Nation 2001: 177-178). Writing quality is of such relevance that even cross-curricular academic achievement is thought to be negatively affected (Santos 1988: 80; Agustín LLach 2007: 2). Throughout

the last three decades, lexis has consequently come to the forefront of investigations into learner language, as epitomised by Nation (1990, 2001), among others.

It is commonplace to assume that with the passing of time students' writing skills will show a degree of lexical improvement more closely resembling the lexical variety associated with native writing. Such a trend is considered to stem from explicit teaching and exposure to the lexical patterns of the foreign language. Specifically, Nation (1981: 60) claims that regular practice in English academic writing should result in an enrichment of students' written output. To date, studies into vocabulary acquisition by non-native learners of English have often assessed the students' productive and receptive mental lexicons (cf. Nation 1990: 11-27; Schmitt 2008: 330-333). The evidence indicates that active vocabulary is acquired at a slower rate than passive vocabulary, as well as to a smaller extent (Laufer and Paribakht 1998; Schmitt 2008: 335). Vocabulary use in the receptive communicative skills (namely reading and listening) is therefore more amenable to a quick progression than that to be found in the productive skills (namely speaking and writing).

A marked difference between native and non-native writing can be found in the dichotomy between formal and informal language, the former relating to academic registers whilst the latter to conversation. Non-native students rely heavily on conversational language (e.g. discourse markers such as *really*, *of course*, *I think*, *don't*) to structure their ideas coherently (Gilquin, Granger and Paquot 2007: 327; Paquot 2010: 193). Underlying spoken and written language (in their extreme versions) is a distinction between congruency and incongruency in natural language use. The acquisition of the L1 is claimed to proceed along this cline, i.e. from congruent to incongruent (Derewianka 2003; Halliday and Matthiessen 2004: 586-658; Achugar and Colombi 2008). The latter mode (associated with academic prose) is only mastered once the learner has completed all his/her formal schooling. Drawing on the latter references, learners could also be assumed to start producing language output which is more congruent, and thus more oral. As their command improves and becomes closer to native standards, the student is arguably more capable of switching between congruent and incongruent modes as required. The problem is therefore not so much using a vocabulary which is more written or spoken, but not knowing when to use it in appropriate contexts.

A further defining feature of writing quality resides in the size of a subject's vocabulary. Nation (2006: 79) proposes that the amount of vocabulary required for comprehension of written discourse encompasses around 8,000 to 9,000 word families (see 3.2. for definition of *word family*). In that instance, 98% of any general English text is supposedly understood and assumed to be available for production by the student. Academic reading, on the other hand, seems to demand a less restricted scope, with an average of 4,000 families to act as an indication of 95% coverage of text (Nation 2001: 147). Vocabulary range in production and comprehension may thus prove to be register-specific.

Research into lexical coverage is frequently based on previously compiled word lists retrieved from corpora. For the purposes of examining lexical competence, four

types of vocabulary can be distinguished, namely high frequency words, academic vocabulary, technical vocabulary and low frequency words (Nation and Coxhead 2001: 252-257; Nation 2001: 187-216). Such a classification draws on a scale based on different word frequencies. High frequency words, first in the list, constitute the 2,000 most frequent word families of the English language, as contained in the *General Service List* by West (1953) (e.g. *talk, food*). Academic vocabulary, second in the scale, comprises the 570 word families identified in Coxhead's (2000) *Academic Word List* (henceforth, *AWL*). These words are reported to be reasonably frequent in a wide range of academic texts (covering around 8%-10% of running words), hence their discourse-organising function (e.g. *assume, achieve, concept*). Technical lexis, third in the scale, is discipline-specific, and covers up to 5% of running words in texts (e.g. *morpheme, phoneme, subordinate clause*). Finally, low frequency words, at the bottom of the scale, are specific to particular texts.

With the aim of measuring the quantity and quality of the lexis used by students, Laufer and Nation (1995: 308) suggest a fourfold approach, each linked to a specific measure: i) lexical originality, ii) lexical density, iii) lexical sophistication, and iv) lexical variation. To start with, lexical originality takes account of the percentage of words typically employed by an individual writer within a group of writers. Lexical density is the percentage count of lexical items found in a written composition. Lexical sophistication is concerned with the frequency of advanced words in discourse. Finally, lexical variation, also referred to as type-token ratio, entails the relationship between different word types and their repetitions. Drawing on these measures, Laufer and Nation (1995: 308) propose an alternative whereby vocabulary is not studied with regard to the text itself, but more importantly, its use is also compared with that of general English. This tool, the so-called Lexical Frequency Profile (henceforward *LFP*), measures the quality and range of a student's productive written vocabulary in respect of a series of frequency lists, namely the first 1,000 most frequent words, the second 1,000 most frequent words (retrieved from West 1953), a list of academic vocabulary (retrieved from Coxhead 2000), and words not included in any of the previous lists. Laufer and Nation's (1995) *LFP* is a widespread measure, even if it relies on a word list compiled in the 1950s (West 1953).

Of particular relevance to this paper is the study by Laufer (1994), in which the *LFP* of language learners of English is measured from a longitudinal perspective. Her article assesses the extent to which students' writing increases in size and range over an academic year. To accomplish this, Laufer (1994) combined a measure of lexical variation (namely type-token ratio) with the *LFP*. The results revealed that the use of more complex lexis (e.g. words belonging to the *AWL*) rises slightly whilst a minimal decrease is apparent across highly frequent words. As acknowledged by the author (1994: 29), such an improvement is still a long way off native-like lexical usage. In view of this finding, the conclusion reached is that size (measured by type-token ratio) does not necessarily correlate with lexical range (measured by the *LFP*).

## 2.2. Error analysis

The students' performance in writing is dependent not only upon the size and diversity of the vocabulary they produce, but equally on how accurately they use it (Schmitt 2008: 333). In other words, lexical richness cannot be properly assessed without accounting for both correct and incorrect uses of individual words. It is therefore essential to acknowledge the contribution of error analysis to come to a more complete understanding of the difficulties learners encounter when getting to grips with a writing task in the L2. Awareness of such problems may lead to better teaching practice, insofar as teachers will be better prepared to help their students tackle their lexical weaknesses.

The concept of error analysis is closely associated with Corder (1967), who brought valuable insights to the field. Paramount in his definition is the dichotomy between *error* and *mistake*, the former referring to a lack of competence on the part of the language learner, and the latter to inaccuracy in the learner's performance. At a later stage, Lennon (1991) put forward a less formalist and more contextual perspective of what constitutes an error. In his view, error identification cannot always be tied to specific words; rather, it can incorporate the surrounding co-text as well as extralinguistic elements. Furthermore, a two-fold distinction is made between *domain*, namely the degree of context present in order for the error to be recognised, and *extent*, namely the amount of linguistic information the speaker is assumed to draw on to correct the error (Lennon 1991: 191).

Learner corpora have been a useful tool for the study of learner language. Based on Kennedy's (1998: 2) definition, learner corpora are understood as "[a] collection of texts written or spoken by non-native speakers of the language in which the corpora are presented". They are of particular relevance to interlanguage analysis, as they shed light on some of the problematic areas experienced by non-native speakers when they are faced with writing in the foreign language, notably in their use of vocabulary. Back in the late 1960s, Dušková (1969: 35-36), based on the analysis of compositions produced by fifty Czech postgraduate English learners, identified four major error categories, namely syntactic errors, errors in construction and government, errors in the use of prepositions, and lexical errors. With regard to lexical errors, six sub-categories were proposed:

- i) Confusions resulting from formal similarity between word pairs: e.g. *same* vs. *some*, *think* vs. *thing*.
- ii) Confusions based on semantically similar words: e.g. *take* vs. *carry*.
- iii) Confusions of related phrases: e.g. *at first* vs. *first*.
- iv) Confusions based on the existence of more than one possible L2 translation for a given L1 word: e.g. *před* vs. *before*, *ago*.
- v) Influence of German: e.g. *chef/chief*.
- vi) Miscellaneous errors: derivational distortions (e.g. *desolve/solve*) and other nonce mistakes (e.g. *I wrote down to my colleague/I wrote to my colleague*).

A remarkable insight in relation to the quantity of lexical errors stems from the influence of not only the mother tongue, but also of other languages learnt previously or simultaneously. Difficulties specifically concerning the L2 system are another driving factor in students' lexical performance. Laufer (1991) also confirms the relevance of this claim, in connection with the concept of *synforms*, namely words that, due to derivational similarity, lead to confusion. Some examples drawn from her taxonomy are *considerable/considerate*, *numerous/numerical*, *historic/historical*, *passion/compassion*, *affect/effect*, *quite/quiet* and *price/prize*.

A more recent example of a lexical error taxonomy is that of Engber (1995), who, using a reasonably small corpus of first-year undergraduate essays (sixty in total), combines a measure of lexical density and variation with a count of lexical errors. Her classification comprises two error types, those in line with semantic misselection of lexical items and those derived from the internal structure of the lexeme. The findings reveal that no significant correlation is apparent between the quantity of lexical items (e.g. nouns, verbs, adjectives, etc.) and the quality of writing. They also reveal a connection between writing quality and the number of lexical errors, insofar as greater quality follows from a decrease in the number of errors (cf. Laufer and Nation 1995; Nation 2001: 177-178; Hawkey and Barker 2004). These results stand in contrast to those reported by Agustín Llach (2007: 11), where no such relationship is evidenced. A possible explanation for this difference lies in the kind of learner corpus employed, namely compositions by nine-to-ten year-olds. Moreover, the fact that her study focuses exclusively on error counts as opposed to a combination of lexical measures with error analysis may have led to this discrepancy. Agustín Llach's (2007: 8) error taxonomy contemplates four categories, namely misspellings (namely orthographic floutings, e.g. *ugli*), borrowings (namely, use of an L1 word, e.g. My sister is *fea*), coinage (namely use of an L1 word with L2 spelling or morphology, e.g. My brother is *agobiated*) and calque (namely literal translation based on L1 patterns, e.g. My *cat black*). Data from the study show a marked dominance of spelling errors (74.8%). This is argued to be unsurprising in view of the limited knowledge that primary school students possess of the English graphological system (Agustín Llach 2007: 11; see also Berman and Verhoeven 2002: 19-21, cited in Bar-Ilan and Berman 2007: 19-21).

This paper is thus intended as an analysis of learner language, with a view to ascertaining what areas prove more problematic for undergraduate EFL productive vocabulary use. The subsequent section describes the procedure followed in order to establish the types of lexical errors committed and, overall, the extent to which the accuracy of students' lexical usage increases over an academic year.

### 3. METHODOLOGY

#### 3.1. *The corpus*

This study draws on *NOSE*<sup>3</sup> (Díaz-Negrillo 2009), an error-annotated learner corpus of English compiled at the Universities of Jaén and Granada (Spain). This learner corpus was collected over a six-year period (2003-2009) as part of a research project

funded by the Spanish Ministry of Education. It comprises a total of 1,173 texts, which amounts to 391,798 words. The rationale behind the creation of this corpus was in the first place the identification of error types by Spanish first-year undergraduates and, subsequently, the design of a fine-grained taxonomy of errors capable of describing accurately and efficiently the errors of Spanish learners of English.

The compositions were produced by first-year undergraduates pursuing the degree in English Studies at the Universities of Jaén and Granada (Spain). Prior to university entry, it is commonplace in Spain for exposure to English in the national curriculum to draw on a grammar-translation style, priming rote learning of grammatical rules and extensive vocabulary lists. This happens at the expense of contextual meaning and real-life interaction in the foreign language. As such, the productive writing ability of Spanish first-year university students is largely underdeveloped.

The compilation process was conducted at three different stages of the academic year, namely beginning (October), mid-way through (February) and at the end (June), during which time the official credit load for this degree totals around 290 hours of tuition in English and 390 hours of tuition in the mother tongue<sup>4</sup>. The texts collected at each of these points correspond to three separate corpus samples, i.e. Sample A (October), Sample B (February) and Sample C (June). The writing task comprises three pre-set argumentative topics and a fourth option of free writing (See Table 1 for details). Students were required to choose one of the four options and write 350 words (maximum) under examination conditions. One hour was allocated to the completion of the task, and the use of additional resources (e.g. dictionaries, grammars, etc.) was forbidden (further details can be found in Díaz-Negrillo 2007).

<b>SAMPLE A</b>	<ol style="list-style-type: none"> <li>1. <i>The importance of foreign languages nowadays.</i></li> <li>2. <i>Getting economically independent.</i></li> <li>3. <i>The internet in society: progress or regress?</i></li> <li>4. Free writing example: <i>A new life.</i></li> </ol>
<b>SAMPLE B</b>	<ol style="list-style-type: none"> <li>1. <i>A destination for a one-week summer holiday: Santo Domingo seaside resort or bohemian and monumental Paris? Give reasons for your choice.</i></li> <li>2. <i>Justify your position in favour or against the following: It's not just that we are slaves of mobile phones but, on top of that, they may be harmful to health.</i></li> <li>3. <i>Express your opinion on the topic of current food diversification: Genetically modified food, organic food, convenient products (fast food, frozen food, etc.)</i></li> <li>4. Free writing example: <i>Does the love exist?</i></li> </ol>
<b>SAMPLE C</b>	<ol style="list-style-type: none"> <li>1. <i>Terrorism in our society.</i></li> <li>2. <i>Do you feel influenced by other cultures' stereotypes? Are they always true? How should we deal with them?</i></li> <li>3. <i>Does your country offer enough job opportunities for a future career? Would you rather move abroad? If so, what are the advantages and disadvantages?</i></li> <li>4. Free writing example: <i>Spain, a tourist resort.</i></li> </ol>

Table 1. Topics

The texts were subsequently transcribed electronically with a view to enabling computerised manipulation of the data. At a later stage, an error annotation scheme (*EARS*, Díaz-Negrillo 2009) was devised and applied to the corpus. The *EARS* error annotation scheme consists of six different error categories, as summarised below:

- i) *Clause grammar (CG)*: Errors in clausal constituents (e.g. subject and predicator), syntactic processes (e.g. coordination and subordination) and multiple structural errors.
- ii) *Phrase grammar (PG)*: Errors in phrasal constituents (e.g. complement and determiner), syntactic processes (e.g. coordination) and multiple structural errors.
- iii) *Word grammar (WG)*: Errors in grammatical categories of word classes (e.g. number, case and tense).
- iv) *Lexis (LX)*: Errors in the use and derivation of lexical units (e.g. nouns, verbs, etc.), as well as in the formation and use of set phrases and collocations.
- v) *Spelling (SP)*: Errors in word lettering, capitalisation and word boundary.
- vi) *Punctuation (PN)*: Errors in punctuation.

Examples of the above can be found in Table 2 below. The first column shows the scope of the annotated unit, the second the tag itself and the third the tag gloss.

Example	Tag	Tag gloss
(1) The leaders think that is better <i>someone who can speak languages</i> (GR-A-1-031-1)	<CG.CS.SB.ON.OR>	Clause grammar, Constituent, Subject, Occurrence, Order
(2) [...] have two legs for <i>walk</i> (GR-A-1-031-1)	<PG.CS.CP.RP.RE.NF.MS>	Phrase grammar, Constituent, Complement, Prepositional Phrase, Realisation, Non-Finite, Misselection
(3) Yesterday I <i>get</i> up at 7:40 (GR-A-1-034-10)	<WG.VR.TN.RT.ER.MS>	Word grammar, Verb, Tense, Remote, External, Misselection
(4) <i>obsseious</i> (GR-A-1-128-10)	<LX.AD.IT.DR.SF.MS>	Lexis, Adjective, Internal, Derivation, Suffixation, Misselection
(5) <i>confortable</i> (GR-A-1-128-10)	<SP.GR.SN.CN.GL.MS>	Spelling, Grapheme, Single, Consonant, General Lettering, Misselection
(6) I will write about what I did yesterday, Why? (GR-A-1-034-10)	<PN.CM.MS>	Punctuation, Comma, Misselection

Table 2. Error categories



The range of error tags within this annotation scheme ranges from the six general categories previously outlined to 612 when providing the maximum amount of detail, as illustrated in column 2 above.

### 3.2. *Data processing*

This study takes its data from the first year of the corpus collection process (2003-2004). This initial batch comprises 179 texts, amounting to 39,015 words, of which 95 correspond to sample A and 49 to sample C. However, in this instance, a selection of just 60 texts (12,500 words) were analysed. Given that one of the goals is to assess the evolution of student's accurate and inaccurate vocabulary use over an academic year, it was necessary, for the purposes of consistency, to narrow down the experimental group to include those participants who completed the task both at the beginning and at the end of the academic year, resulting in 30 subjects. Sample B was discarded on the grounds that more progression is likely to ensue from a longer period of time. There is also an implication that after exposure to the complete credit load in English for the first year, students' productive writing skills may show some degree of improvement.

Following the retrieval of the study sample, a number of issues were addressed, most of which emerged throughout the investigation process. The initial idea was to measure the complexity of the students' lexical range over a period of an academic year. With this in mind, the Lexical Frequency Profile (Laufer and Nation 1995: 308) was deemed the most convenient measure for the immediate purposes of this paper. For the sake of accuracy and consistency, the use of such software as WordSmith Tools (Version 5) (Scott 2008) or Paul Nation's RANGE32 (<http://www.vuw.ac.nz/lals/staff/paul-nation/RANGE32.zip>) was discarded. RANGE32 is an ICT tool designed to run the lists within the Lexical Frequency Profile on any given text. Considering that this paper draws on data from sixty texts, RANGE proved to be limited in its scope, allowing only thirty-two texts at a time to be fed into the programme. Additionally, RANGE poses problems for the analysis of learner language, due to the high number of spelling errors (e.g. *negociate*, *reasure*, *stablish*), which are not identified by the programme. A preliminary trial using the dataset showed that 30% (97 tokens) of the words pertaining to sample A (namely thirty texts) were spelling errors which were consequently misclassified. Although just the wrongly classified items could have been sifted through manually, a completely manual analysis was deemed most appropriate in view of the problems involved in an automatic corpus-assisted analysis and for the purposes of efficiency.

Further to the issues raised so far, mention should be made that RANGE32 appears in two versions: i) the one employed by Laufer (1994) (see section 2.1) and ii) a corpus-based set of sixteen frequency lists taken from the *British National Corpus* (BNC henceforward). Whilst the first version is clearly more manageable due to its size, the general vocabulary frequency list employed in this case dates back to the 1950s, and so is arguably outdated for current use. As regards the AWL, several studies have highlighted a limited coverage of discipline-specific academic English (Chung and Nation 2004;

Hyland and Tse 2007). The second and more comprehensive version of RANGE32, which was used in this paper, despite its more extensive collection of vocabulary, still fails to include words that are frequently used nowadays, such as *e-mail*, *internet*, *mobiles*. This may be explained by reference to the date of compilation of the corpus (early nineties), when such technological developments were not as well-known as nowadays. This situation was addressed in this paper through the creation of two supplementary lists, namely list 17 and list 18. The former was designed to cater for those items not present in any of the lists, but which either exist in English (e.g. *videoconferences*, *websites*) or for which there is no direct translation (namely Spanish proper nouns, like Christian names or toponyms: *Pili*, *Jaen*). The latter, on the other hand, entails non-existing words used by learners, similar to Agustín Llach's (2007) categories of borrowing and coinage of L1 items (see section 2.2. above) (e.g. *aproveded=took advantage of*, *bonobus=travel card*, *electrodomestic=appliance*). Instances of misspellings or derivational errors for which understanding of the intended item is not impeded were assigned to the BNC list where the correct version occurs. Accordingly, the derivationally malformed adjective *\*affectionable* is annotated as falling within frequency list 3, where the intended correct version *affectionate* occurs. Similarly, *\*consequences*, an example of a misspelling, is treated as belonging to frequency list 4, in spite of its erroneous form.

The frequency lists mentioned above rely on the concepts of *word family* and *type*. The term *word family*, as put forward by Bauer and Nation (1993: 253), includes base forms, in addition to their inflectional (e.g. third person singular present simple tense *-s*) and derivational variants (e.g. *develop* vs. *development*). To provide an example, the items *accuse*, *accusation*, *accusations* and *accusingly* would qualify as *types* within one and the same word family, namely *accuse*. With this in mind, a *type* then would correspond to each different lexical unit, be it a derivative or a word form. The application of these notions to the analysis of data may be called into question on the grounds that the attention given to meaning and word-class appears to be minimal. To provide an example, one cannot tell whether homonymous pairs such as *will* and *may* are being treated as modal verbs or nouns. Similarly, with the case of *that*, no distinction is apparent between its use as a subordinator, a demonstrative pronoun or a determinative.

For the sake of applying the aforementioned measure, the two samples were converted into two separate word lists. Subsequently, the two lists were manually annotated based on the family and type categorisation provided by the sixteen BNC frequency lists. Each item was therefore assigned two different pieces of information, one corresponding to the specific word list it belongs to, and the other accounting for whether the item qualifies as a word family. For instance, the word *access* is classified as belonging to list 2, this representing also a word family. The results drawn from this analysis are displayed in section 4.1, where only types are shown, on account of the close similarity obtained from the count of word families and types. The next step involved the identification of the number of correct and incorrect tokens in both samples, with the intention of establishing whether any development in the accuracy of students' lexical

competence occurs over an academic year (cf. section 4.2). Correct tokens here refer to every correct occurrence of a lexical item, whilst incorrect tokens concern each time an error tag is applied to a lexical item. Therefore, in the instance that a word is tagged more than once, the quantification of incorrect tokens corresponds to the number of tags. An example is provided in (7). Error annotation for this item applies two tags; the first accounts for a lexical mistake (i.e. incorrect derivation), and the second involves a word grammar mistake (i.e. singular instead of plural agreement).

- (7) *there are many* <LX.NN.IT.DR.BL.MS><WG.NN.NB.SG.ER.MS> *electro-domestic* </WG.NN.NB.SG.ER.MS></LX.NN.IT.DR.BL.MS> (GR-A-1-034-10)

To narrow the scope, incorrect tokens were further analysed in terms of the six error categories previously described. The rationale behind the concept of error in this paper contemplates not only cases of derivationally and semantically misselected items (e.g. Engber 1995, section 2.2), but also the erroneous combination of lexical items at clause and phrase level. Lexical competence thus constitutes knowledge of the form and meaning of the word and, of equal importance, where, when and how to use the word appropriately (Nation 2001: 27).

Each of the six error categories was investigated in search of the number of incorrect tokens pertaining to each one. In addition to the study of errors at their most basic level, the analysis delved further into two sub-categories commonly present in many of the error tags, namely internal and external errors. Such a distinction (Díaz-Negrillo 2009: 21) is inspired, among others (e.g. Granger and Monfort 1994; Flowerdew 2006), by Köhlmeyer's (2003: 31) twofold classification into "category errors" and "realisation errors". The former can be defined as those cases in which a word on its own is malformed (see 8, featuring an extra derivational suffix, i.e. *provoke*> *provocate*), whereas the latter contains words that are correctly formed, but are not used in the correct context. This is the case in example 9, where the inflectional form employed (i.e. singular) does not agree with that of the demonstrative determiner (i.e. plural).

- (8) <LX.VR.IT.DR.SF.OV>*provocate*</LX.VR.IT.DR.SF.OV> (GR-A-1-053-3)

- (9) [...] basic notions of *these* <WG.NN.NB.SG.ER.MS>*language* </WG.NN.NB.SG.ER.MS> (GR-A-1-022-1)

#### 4. RESULTS AND DISCUSSION

This section presents and discusses the results obtained in this study. It is laid out in such a way as to address each of the steps taken throughout the analysis of the dataset. As such, 4.1. offers the findings stemming from the application of the LFP. The following section (4.2) provides evidence gleaned from the study of errors. Findings are in all cases expressed in percentages.

4.1. *Lexical Frequency Profile*

Table 3 below brings together the outcome of the LFP in both sample 1 (S1) and sample 3 (S3) across the eighteen lists considered.

BNC lists	Types (S1)	Types (S3)
1	66.19	68.21
2	16.94	17.41
3	5.55	4
4	2.54	3.2
5	1.41	1.33
6	0.84	0.71
7	0.65	0.98
8	0	0.18
9	0.47	0.36
10	0.094	0.18
11	0.18	0
12	0.094	0
13	0.094	0
14	0.094	0.27
15	2.63	1.33
16	0	0
17	1.60	1.6
18	0.56	0.27

Table 3. Comparison of types in both samples.

As evidenced above, those highly frequent words belonging to lists 1 and 2 clearly stand out. This would appear to imply that students' vocabulary range is as yet quite limited, resulting from an over-reliance on the most common English words (e.g. *people*, *important*), at the expense of semantically more complex and therefore more advanced lexical items (e.g. *importance*, *moreover*). A negligible increase over an academic year is found however in lists 2, 4 and 7, which would indicate that there are exceptions to the aforementioned general tendency. The increase observed in list 1 is evidence of students' persistent use of the most frequent words. A possible explanation here is the limited time span considered in this study, inasmuch as students seem to require more time and exposure to the language before their vocabulary range undergoes a significant improvement. The results partly confirm Laufer (1994), who, notwithstanding a different methodological approach (namely three frequency lists), reports similar findings (see 2.1).

The students' predisposition to employ such a general lexis is perhaps echoed by Paquot's (2010: 145) claim that EFL students revert to using vaguer and more general words, as a result of a lack of register awareness. EFL students, unlike the average native speaker, are thought to be more drawn to an informal, oral-like style in their written assignments. Furthermore, due to a lack of vocabulary, learners are often forced to opt for familiar nouns which they frequently repeat (Hinkel 2004: 96). Consequently, and in line with Hasselgren (1994), intermediate EFL learners would appear to reach a plateau where their use of vocabulary at that point struggles to move forward.

#### 4.2. *Error analysis*

As mentioned in 1 and 3.2, the next step involved looking at the extent of correct and incorrect usage of vocabulary in both samples. Figures 1 and 2 below contain the same data, the only difference being the inclusion or exclusion of Clause grammar multiple structural errors (henceforward, CG.ME). These errors consist in instances where the allocation of error tags is problematic on the grounds of multiple errors occurring within a single clause, to the extent that no right version can be retrieved and the resulting text becomes incomprehensible. An example of a string of words annotated as CG.ME is the following:

- (10) <CG.ME>The fact you find a job or not even being in our country or in a foreing country is linked to the urge of progress, to your feelings of working, to better but on the other hand, luck is there and to keep a good friends and contacts can help you so it's important</CG.ME>

Figure 1 counts every word included within the CG.ME tag as a separate incorrect token, while figure 2 excludes these error types from the calculations. This said, figure 1 illustrates a decrease in the percentage of correct tokens and, in turn, an increase in the percentage of incorrect tokens. This paper examines the statistical relevance of these figures according to Pearson's chi-square test. This test can be used for the significance of occurrences and tells whether given frequencies of occurrence are random or motivated. Yates's correction was then used for further certainty that the significance of the values obtained is not an overestimation of low data in the dataset. The correction is considered a conservative revision of samples where, as in this case, the *n* value is occasionally low. This revision compensates the low *n* with a higher relevance threshold. In the figures at issue, the difference was found to be statistically significant, in that there is a noticeable relation between the sample and the percentage of correct/incorrect cases (chi Pearson: 7.89, *p*=0.0049; 7.75, *p*=0.0054 after Yates correction). By contrast, in figure 2, an increase in correct tokens and a decrease in incorrect tokens can be seen. The difference here proves to be non-significant (chi Pearson: 0.06, *p*=0.81; 0.05, *p*=0.082 after Yates correction).

These two figures are an indication of the influence that different error counts for a specific error category can have upon the results obtained. This confirms the difficulty

experienced by the error analyst when deciding on the scope of the error tag. As such, error identification at times requires a method of quantification that fulfils the purposes of the investigation at hand. In what follows, CG.ME errors are removed from the total error count, in view of the possible skewing that might result from counting each individual token within this tag.

Given that students' overall use of vocabulary does not gain in accuracy over an academic year, it was deemed appropriate to further delve into the six error categories comprising the *EARS* error taxonomy. Figure 3 below displays the percentages across the six error levels and across both samples.

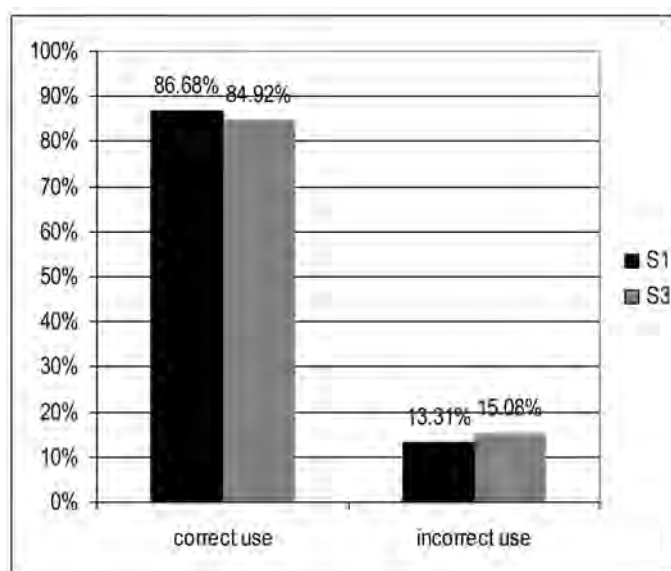


Figure 1. Correct and incorrect tokens in both samples (with CG.ME).

ACCURACIES AND INACCURACIES IN ELF LEARNERS' WRITTEN VOCABULARY USE

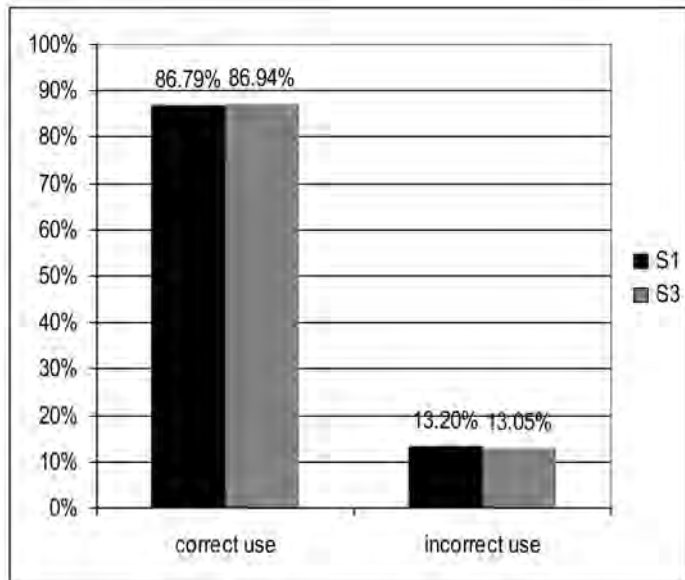


Figure 2. Correct and incorrect tokens in both samples (without CG.ME).

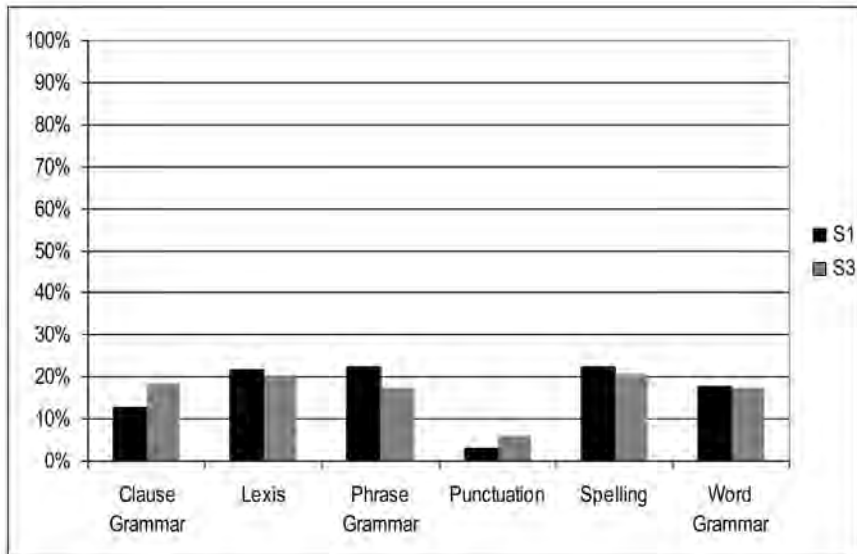


Figure 3. Incorrect tokens sorted by error type.

An overall decrease in error types is evident in all categories, except for Clause grammar and Punctuation, where the opposite applies. There is a significant relation

between the sample and the error categories studied here (chi Pearson: 21.92,  $df=5$ ,  $p=0.001$ ). This means that the differences in error frequency recorded for each category are conditioned by the sample in each case. Thus, there is a significant increase in errors concerning Clause grammar and Punctuation between the first and the third sample of the academic year. By contrast, there is a significant decrease in errors concerning Lexis, Phrase grammar, Spelling and Word grammar between the first and the third sample of the academic year.

One possible interpretation of the above findings suggests how students' awareness of the linguistic system increases with regard to word level categories (Lexis, Spelling and Word grammar), but not in relation to their function in discourse. Considering the short time span of the sample, it is arguably no surprise that students' interlanguage progresses at such a slow rate. Such a tendency may be indicative of the fact that language learners advance at a quicker pace at those language levels coming closer to the bottom of the linguistic rankscale (namely word and phrase level), but struggle with more complex levels, such as clause, sentence and paragraph. This pattern is also seen to apply to the linguistic development inherent in the native acquisition of the language, as native nine-to-ten year olds also experience difficulty with both the orthographical system, as well as the appropriate use of punctuation (Berman and Verhoeven 2002: 19-21, cited in Bar-Ilan and Berman 2007: 27).

In order to test the idea that the use of words in longer stretches of discourse poses more difficulties to students than their context-independent use, attention was subsequently paid to the dichotomy between internal vs. external errors (see section 3.2). Figure 4 presents the outcome of such a count.

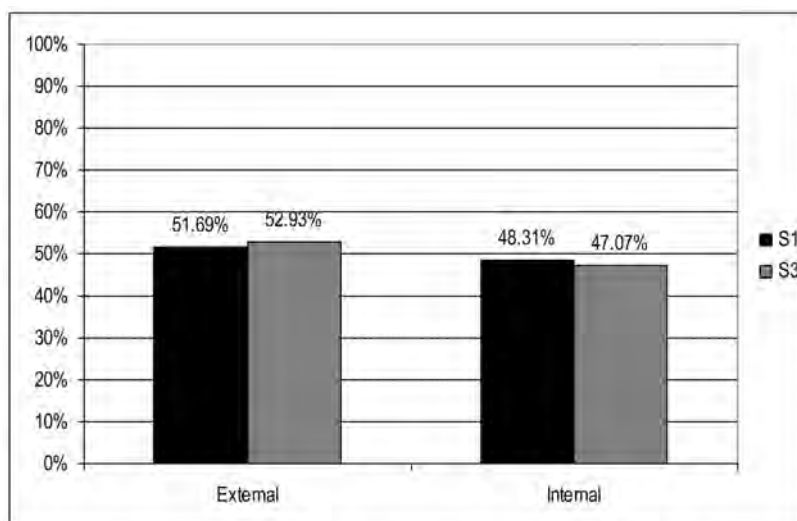


Figure 4. External and internal errors.



The percentages indicate that external errors marginally outweigh internal errors. This appears to be a tendency common to most EFL learners, in that knowledge of the internal make-up of words seems to prevail over the meaning and contextual features surrounding the cohesive use of lexis in discourse. Flowerdew (2006) seems to support this finding, contending that EFL learners often misuse signalling nouns such as *factor* or *issue*, those of which endow discourse with cohesion and coherence. Furthermore, Paquot (2010: 148) highlights that EFL learners overuse certain word classes (adverbs, prepositions and conjunctions) to ensure cohesion. When comparing the results for both samples, the statistical test highlights a non-significant relation between sample and error types (chi Pearson: 0.14,  $p=0.71$ ; 0.09,  $p= 0.76$  after Yates correction). Consequently, once again the short time span considered is thought to prevent the occurrence of a more marked improvement.

## 5. CONCLUSIONS

This paper could have implications for the teaching and learning of vocabulary by EFL learners. First and foremost, it is evident that one of the areas students have difficulty with is that of vocabulary. Lexical knowledge has been shown to prove fundamental for the development of the passive communicative skills (i.e. reading and listening) (e.g. Saville-Troike 1984). Productive use of vocabulary is more complex, because it relies on its prior acquisition through the receptive skills. In fact, as pointed out by Nation (2001: 182), improving vocabulary knowledge is not as easy a task as using it accurately in appropriate productive settings (be it writing or speaking).

As evidenced in this paper (especially section 4.1.), students' interlanguage is often dependent on a repetitive use of vocabulary. Their vocabulary range tends to draw on words that, whilst being very frequent, lack semantic and morphological sophistication. Support for this idea can be found in Nation (1990: 22), who contends that students tend to opt for words that are shorter in length and simpler, (e.g. *thing*), as opposed to more morphologically complex words (e.g. *accusation*). This situation might be tackled by fostering students' awareness of word-building devices, insofar as this may also contribute to an enhancement of their autonomous learning skills. Learners, ideally, should wean themselves off the dictionary and the teacher through the increasing establishment of connections among word-family members (Schmitt and Zimmerman 2002: 147).

It has also been implied in this study that, in addition to examining vocabulary range per se, one cannot overlook the degree of precision and correctness in the lexical component of EFL writing. The teacher needs to be well informed about the errors students commit in order to implement remedial techniques that assist students in avoiding these in the future. Error analysis, however, is by no means free of complications, as certain decisions need to be taken that will impact on the outcome of such an analysis. Paramount among such decisions is the scope of the error tags employed in the error count. In learner language, long stretches of language may be annotated as either one major structural error or, alternatively, as a sequence of specific

errors. This is ultimately not only according to the error annotation system used or the accuracy with which it is used, but also according to the language description level: this type of errors are thus likelier to occur where long sequences are involved (e.g. Clause grammar compared with Word grammar).

The longitudinal approach of this paper has revealed only slight progression in certain areas, namely Lexis, Phrase grammar, Spelling and Word grammar. For those areas where students clearly struggle, namely Clause grammar and Punctuation, emphasis should be laid on the teaching and learning of vocabulary in context. As such, the development of lexical competence should never rest on the sole memorisation of words, but more importantly, incorporate information about their typical patterning and collocations. Only by taking account of the overall syntactic environment, will one be better equipped to lead students to overcome Clause grammar and Punctuation errors. Corroboration for this may be found in Laufer and Sim (1985), where the conclusion reached is that EFL learners should prioritise their linguistic knowledge in the following order: vocabulary, subject matter and then syntactic structure. Whilst no exact reasons for the findings in this study can be provided, certain variables, such as the group surveyed or the teaching process used (university compared with secondary education) may have a bearing on the results obtained.

Possible future avenues for further research would contemplate a more specific approach in terms of the error count, paying attention not only to the six general error categories and the two sub-levels of internal vs. external, but to possible associations between all other sub-levels (cf. Díaz-Negrillo and Valera 2010). In addition, it is important to recognise that no generalisations should be made on such a small sample. Consequently, the replication of the results obtained needs to be tested on a larger sample. Finally, on account of the absence of certain current and highly common words (e.g. *internet*, *e-mail*) from the frequency lists employed (BNC), a possible alternative would be to make use of frequency lists drawn from a more updated corpus, such as the Bank of English.

## ACKNOWLEDGEMENTS

Our warmest thanks to Lynette and John Bartley, Ana Díaz-Negrillo, Encarnación Hidalgo Tenorio and Salvador Valera Hernández.

## NOTES

\* Correspondence to: Leanne Bartley. Dept. de Filologías Inglesa y Alemana. Facultad de Filosofía y Letras. Campus de Cartuja, s/n, 1807. Granada, España. E-mail: lbartley@ugr.es.

1. This paper has been funded by the Spanish Ministry of Education (research project HUM2007-60107/FILO and research grant FPU AP2008-03204).
2. **Error Annotation Retrieval Scheme**
3. **Non-native Spanish Corpus of English**
4. Tuition in English comprises three compulsory modules on English linguistics, English literature and Communicative Skills (productive and receptive).

## REFERENCES

- Achugar, M. and M. C. Colombi. 2008. "Systemic functional linguistic explorations into the longitudinal study of advanced capacities: The case of Spanish heritage language learners". *The Longitudinal Study of Advanced L2 Capacities*. Eds. L. Ortega and H. Byrnes. New York and London: Routledge. 36-57.
- Agustín Llach, M. P. 2007. "Lexical errors as writing quality predictors". *Studia Linguistica* 61(1): 1-19.
- Bar-Illan, L. and R. A. Berman. 2007. "Developing register differentiation: The Latinate-Germanic divide in English". *Linguistics* 45(1): 1-35.
- Bauer, L. and P. Nation. 1993. "Word families". *International Journal of Lexicography* 6(4): 253-279.
- Berman, R. A. and L. Verhoeven. 2002. "Cross-linguistic perspectives on the development of text production abilities: speech and writing". *Written Language and Literacy* 5(1): 1-43.
- Chung, T. M. and I. S. P. Nation. 2004. "Identifying technical vocabulary". *System* 32: 251-263.
- Corder, S. P. 1967. "The significance of learners' errors". *International Review of Applied Linguistics* 5(4): 161-170.
- Coxhead, A. 2000. "A new academic word list". *TESOL Quarterly* 34(2): 213-238.
- Derewianka, B. 2003. "Grammatical metaphor in the transition to adolescence". *Grammatical Metaphor: Views from Systemic Functional Linguistics*. Eds. A. M. Simon-Vandenberg and M. Taverniers. Amsterdam/Philadelphia: John Benjamins. 185-219.
- Díaz-Negrillo, A. 2007. *A Fine-Grained Error Tagger for Learner Corpora*. Unpublished Doctoral Dissertation. Universidad de Jaén.
- Díaz-Negrillo, A. 2009. *EARS. A User's Manual*. Munich: Lincom.
- Díaz-Negrillo, A. and S. Valera. 2010. "A learner corpus-based study on error associations". *Procedia Social and Behavioural Sciences* 3: 72-82.
- Dušková, L. 1969. "On sources of errors in foreign language learning". *International Review of Applied Linguistics in Language Teaching* 7(1): 11-36.
- Ellis, R. and G. Barkhuizen. 2005. *Analysing Learner Language*. Oxford: Oxford University Press.
- Engber, C. A. 1995. "The relationship of lexical proficiency to the quality of ESL compositions". *Journal of Second Language Writing* 4(2): 139-155.
- Ferris, D. R. 2001. "Teaching writing for academic purposes". *Research Perspectives on English for Academic Purposes*. Eds. J. Flowerdew and M. Peacock. Cambridge: Cambridge University Press. 298-314.
- Flowerdew, J. 2006. "Use of signalling nouns in a learner corpus". *International Journal of Corpus Linguistics* 11(3): 345-362.
- Gilquin, G., S. Granger and M. Paquot. 2007. "Learner corpora: The missing link in EAP pedagogy". *Journal of English for Academic Purposes* 6: 319-335.

- Granger, S. 1998. *Learner English on Computer*. London: Addison Wesley Longman.
- Granger, S. and G. Monfort. 1994. "La description de la compétence lexical en langue étrangère: perspectives méthodologiques". *Acquisition et Interaction en Langue Etrangère (AILE)* 3: 55-75.
- Halliday, M.A.K and C. Matthiessen. 2004. *An Introduction to Functional Grammar*. London: Hodder Education.
- Hasselgren, A. 1994. "Lexical teddy bears and advanced learners: a study into the ways Norwegian students cope with English vocabulary". *International Journal of Applied Linguistics* 4(2): 237-258.
- Hawkey, R. and F. Barker. 2004. "Developing a common scale for the assessment of writing". *Assessing Writing* 9: 122-159.
- Hinkel, E. 2002. *Second Language Writers' Text: Linguistic and Rhetorical Features*. Mahwah, N.J.: Lawrence Erlbaum.
- Hinkel, E. 2004. *Teaching Academic ESL Writing: Practical Techniques in Vocabulary and Grammar*. London and Mahwah, N.J.: Lawrence Erlbaum.
- Hyland, K. and P. Tse. 2007. "Is there an 'Academic Vocabulary'?" *TESOL Quarterly* 41(2): 235-253.
- James, C. 1998. *Exploring Error Analysis: Errors in Language Learning and Use*. London: Longman.
- Kennedy, G. 1998. *An Introduction to Corpus Linguistics*. London and New York, NY: Longman.
- Köhlmyr, P. 2003. "To Err is Human..." *An Investigation of Grammatical Errors in Swedish 16-year-old Learners' Written Production in English*. Göteborg: Acta Universitatis Gothoburgensis.
- Laufer, B. 1986. "Possible changes in attitude towards vocabulary acquisition research". *International Review of Applied Linguistics in Language Teaching* 24(1): 69-75.
- Laufer, B. 1991. "Some properties of the foreign language learner's lexicon as evidenced by lexical confusions". *International Review of Applied Linguistics in Language Teaching* 29: 317-330.
- Laufer, B. 1994. "The lexical profile of second language writing: does it change over time?" *RELC Journal* 25: 21-33.
- Laufer, B. and I. S. P. Nation. 1995. "Vocabulary size and use: lexical richness in L2 written production". *Applied Linguistics* 16: 307-322.
- Laufer, B. and T. S. Paribakht. 1998. "Relationship between passive and active vocabularies: Effects of language learning context". *Language Learning* 48: 365-391.
- Laufer, B. and D. D. Sim. 1985. "Taking the easy way out: non-use and misuse of contextual clues in EFL reading comprehension". *English Teaching Forum* 23(2): 7-10, 20.
- Lennon, P. 1991. "Error: Some problems of definition, identification and distinction". *Applied Linguistics* 12(2): 180-196.
- Nation, I. S. P. 1981. "Free composition and remedial work". *RELC* 6: 56-68.
- Nation, I. S. P. 1990. *Teaching and Learning Vocabulary*. Boston: Heinle and Heinle.

- Nation, I. S. P. 2001. *Learning Vocabulary in Another Language*. Cambridge: Cambridge University Press.
- Nation, I. S. P. 2006. "How large a vocabulary is needed for reading and listening?". *Canadian Modern Language Review* 63(1): 59-82.
- Nation, I. S. P. and A. Coxhead. 2001. "The specialised vocabulary of English for academic purposes". *Research Perspectives on English for Academic Purposes*. Eds. J. Flowerdew and M. Peacock. Cambridge: Cambridge University Press. 252-267.
- Paquot, M. 2010. *Academic Vocabulary in Learner Writing: From Extraction to Analysis*. London and New York: Continuum.
- Santos, T. 1988. "Professors' reactions to the academic writing of nonnative-speaking students". *TESOL Quarterly* 22(1): 69-90.
- Saville-Troike, M. 1984. "What really matters in second language learning for academic achievement?" *TESOL Quarterly* 18: 199-219.
- Schmitt, N. 2008. "Review article: Instructed second language vocabulary learning". *Language Teaching Research* 12(3): 329-363.
- Schmitt, N. and C. B. Zimmerman. 2002. "Derivative word forms: What do learners know?" *TESOL Quarterly* 36(2): 145-171.
- Scott, M. 2008. *WordSmith Tools, Version 5*. Liverpool: Lexical Analysis Software.
- Silva, T. 1993. "Toward an understanding of the distinct nature of L2 writing: The ESL research and its implications." *TESOL Quarterly* 28: 657-677.
- West, M. 1953. *A General Service List of English Words*. London: Longman.

