

## Contrastive Study of Lexical Profiles of International and U.S. Lectures Delivered in English

MILICA VUKOVIĆ-STAMATOVIĆ

Univerzitet Crne Gore

vmilica@ucg.ac.me

International academic contexts where English is used as a *lingua franca* (ELF) have become ubiquitous. ELF lectures have been studied from a number of perspectives, but they have not been lexically profiled. We depart from the assumption that the lexical profile of academic lectures delivered in international settings may differ from that of lectures delivered in Anglophone contexts, and that these differences have pedagogical implications for the teaching and learning of academic English from an ELF-perspective. We lexically profile a corpus of fifty university lectures delivered in English in five European countries and compare them against sixty-two lectures delivered in English in the U.S. We find that 3,000 words are needed for good listening comprehension in both sets of lectures, while ideal comprehension is reached at 11,000 words for international and 7,000 words for U.S. lectures, which suggests differences between the two in terms of variation in low-frequency vocabulary. Some function words are much more frequent in international than in U.S. lectures. International lectures also feature less high-frequency and more mid-frequency academic vocabulary than U.S. lectures. These differences mostly reflect the use of ELF-specific communicative strategies in international lectures. Focusing on them and potentially making academic ELF-specific word lists may ensure the more efficient teaching of academic English from an ELF-perspective.

Keywords: ELF; EAP; academic lectures; lexical profile; academic vocabulary

...

## Un estudio contrastivo de los perfiles léxicos de clases magistrales en inglés en universidades internacionales y de EEUU

Los contextos académicos internacionales en los que se usa el inglés como lengua franca se pueden encontrar en ámbitos muy diversos. Hasta ahora se han estudiado las clases magistrales que se imparten en esta variedad de la lengua desde perspectivas igualmente diversas, pero no se ha estudiado su perfil léxico. Partimos de la base de que el perfil léxico de las clases magistrales que se imparten en contextos internacionales puede ser distinto del existente en contextos anglófonos, y de que estas diferencias afectan, en el plano pedagógico, a la enseñanza y aprendizaje del inglés académico desde la perspectiva del inglés como lengua franca. En este artículo se ha estudiado el perfil léxico de cincuenta clases magistrales universitarias impartidas en inglés en cinco países europeos diferentes y se han comparado con sesenta y dos clases magistrales en inglés impartidas en EEUU. Hemos descubierto que se necesitan 3.000 palabras para conseguir una buena comprensión oral en ambos grupos de clases, al mismo tiempo que se consigue una comprensión oral ideal con 11.000 palabras en las clases internacionales y 7.000 en las que se desarrollaron en EEUU. Este dato da a entender que hay diferencias notables entre ambos contextos en cuanto a las variaciones en el vocabulario de baja frecuencia. Algunas palabras de tipo funcional son mucho más habituales en las clases internacionales que en las que se imparten en EEUU. En las clases internacionales también se usa menos que en las clases de EEUU el vocabulario académico de mayor frecuencia y más el vocabulario de frecuencia media. Todas estas diferencias están relacionadas con el uso de estrategias comunicativas propias del inglés como lengua franca en las clases internacionales. Un enfoque más centrado en estas estrategias y en el uso de listas de palabras propias de esta variedad puede dar como resultado un proceso de enseñanza del inglés académico más eficiente desde una perspectiva vinculada al inglés como lengua franca.

Palabras clave: Inglés como lengua franca; clases magistrales; perfil léxico; vocabulario académico

### 1. INTRODUCTION

Nowadays, a large number of international students attend university lectures delivered in English as a Lingua Franca (ELF), i.e., in English as a common means of communication for teacher(s) and students, some of whom have different mother tongues. Although these are typical ELF university settings, ELF in fact refers to any settings when English is spoken between speakers where at least one of them uses it as a second language (Mauranen 2018, 8). Other ELF definitions offer a broader understanding of the term—thus, for Jenkins (2015, 73), it refers to any “multilingual communication in which English is available as a contact language of choice, but is not necessarily chosen” and we could actually speak of English as *Multilingua* Franca (EMF).

What such a definition addresses, among other things, is that other languages might be present in this contact communication, either explicitly (for instance, through code-switching) or implicitly (given that the first language of the speakers will always exert some influence on this communication).

While research on the lexical profile of academic English delivered in some Anglophone contexts (typically by native English lecturers for a prevalingly Anglophone student audience) has been carried out (Dang and Webb 2014; Coxhead et al. 2017; Dang 2018), we should not automatically assume that the lexical make-up of such lectures will be the same as those delivered in English in international settings (typically by multilingual lecturers for a predominantly international student audience; henceforth: international lectures in English). Amongst the factors which may have a bearing on the vocabulary profile of the delivery are the languages that the lecturer speaks, including not just what their first language is but also whether they speak any other additional language(s) besides English, as well as lecture settings themselves, given that lecturer's accommodation to their audience is "at least equally relevant to explaining what happens in successful ELF communication," as Mauranen argues (2018, 12). Thus, both native and non-native speakers of English may use creative language forms, possibly diverging from standard English, in order to make themselves understood and be efficient in ELF settings (c.f. Ranta 2018). This feature, among others, of ELF enables successful communication. Findings also imply that the effectiveness of speakers in ELF settings depends more on their pragmatic abilities and less on their English proficiency (Björkman 2011, 85).

Due to an unprecedented rise in the internationalization of universities and the accompanying number of courses delivered in English, ELF academic contexts, which display a lot of variation amongst themselves, are ubiquitous worldwide and warrant more linguistic investigation, especially given that this global outreach of English has a number of pedagogic implications (Dewey 2012; Dewey and Pineda 2020). The majority of learners of English as an additional language "do not make anywhere near the same vocabulary gains as native speakers, and subsequently they do not have anywhere near the vocabulary size of native speakers," Nation and Anthony (2017, 366) conclude in their review of vocabulary size studies conducted in a number of Anglophone and non-Anglophone countries. As vocabulary knowledge is a good predictor of comprehension (Laufer and Ravenhorst-Kalovski 2010), there may be justification for the concern as to whether university students speaking English as an additional language will have command of a sufficiently large vocabulary, i.e., know about 95% of the words used in a lecture to ensure comprehension—as per measures made by van Zeeland and Schmitt (2013). Additionally, researchers have argued that lecture comprehension will also depend on students' knowledge of academic vocabulary (Dang and Webb 2014; Dang et al. 2017), which is why it should be determined how much and which academic vocabulary is needed for the purpose of following ELF academic lectures. There is certainly a need to make sure that modern English Language Teaching (ELT)

is “effective in preparing learners to use ELF in global settings” (Galloway 2018, 468). Native speakers of English need to be prepared for ELF communication too, as “nobody is a native of ELF” (Jenkins 2014, 38).

Bearing the above in mind, determining the lexical profile of international lectures in English and whether and how they differ from lectures in English delivered in Anglophone settings could have important pedagogical implications for teaching academic English from an ELF-perspective, and it is this goal that drives the present study. To investigate this issue, we will compare the lexical profile of a corpus of lectures delivered in English in five European countries (ELF contexts) against a corpus of lectures in English from the U.S., as representative of an Anglophone context. In light of the above, we pose the following research questions:

1. Given ELF specificities, is the general lexical profile of international academic lectures different from that of U.S. academic lectures in English?
2. What are the differences between international and U.S. academic lectures in English in terms of their use of high- and mid-frequency words?
3. How much high- and mid-frequency academic vocabulary is present in international- compared to U.S. academic lectures in English?

To answer these questions, we employ the methodology of lexical frequency profiling to determine the vocabulary profile of a corpus of international university lectures in English, as well as that of a corpus of U.S. lectures in English, for comparison purposes.

## 2. THEORETICAL BACKGROUND

Here we discuss the relation between vocabulary size and listening comprehension, after which we review the methodology of lexical frequency profiling and word lists. Finally, we briefly reflect on the vocabulary features of spoken academic English and teaching English from an ELF-perspective.

### 2.1. Vocabulary size and listening comprehension

There is a connection between vocabulary size and level of comprehension, for both native and non-native speakers of a language (Andersen and Freebody 1981; Nagy 1988; Schmitt et al. 2011). More studies have focused on the vocabulary thresholds needed for reading comprehension than on those required for listening comprehension—those dealing with the latter are more recent in character and rely on the findings of the former.

Laufer (1989) determined that a reader should know at least 95% words of an English text to understand it “reasonably,” while Nation (2006b) suggested that a reader should, in fact, know 98% of the words used in a text for its ideal comprehension. The latter typically translates to as many as 8,000-9,000 word families, Nation adds. The same two percentages are quoted for listening comprehension—Van Zeeland and

Schmitt (2012) find that knowledge of 95% of the words are needed for adequate listening comprehension, while a listener should know 98% of the words for optimal listening comprehension. It is obvious that the required vocabulary sizes of thousands of words may be difficult to reach for many international English speakers. In fact, as Milton (2010) and Capel (2012) point out, those with a command of English at the intermediate (B1)<sup>1</sup> level only have about 3,000 word families; upper-intermediate (B2) speakers typically know 4,000 word families; finally, the vocabulary sizes of 5,000 word families or more can be expected of advanced and proficient speakers (C-levels). On the other hand, estimates of vocabulary sizes for young adults whose first language is English (based on research in several Anglophone countries) range from about 10,000 to 17,000 word families (Duff and Brydon 2020, 8).

## 2.2. Lexical profiling

Lexical (frequency) profiling is a method which emerged with the rise of electronic corpora and modern computer programs—it was first introduced by Laufer and Nation in 1995. The lexical profile of a corpus is obtained by loading the corpus under investigation into a specialized lexical profiling software which calculates the coverages of various word lists in that corpus. Thus, we can determine the proportion of the corpus that is covered by, for example, the most frequent 2,000 words of English, proper names, frequent academic words, etc., depending on which word list is chosen for profiling. The lexical profile obtained can then be compared with those established for other corpora, with many research and pedagogical implications—for instance, we may establish which corpora are more lexically demanding than others, which are more suitable to be included in the materials for academic or specialized language learning, etc.

Lexical profiling is a well-known frequency-based measure of vocabulary and is widely used in general vocabulary research. The present study attempts at applying it in ELF research as well. Of course, other methods (for instance, the Type-to-Token Ratio method) can be used to determine the lexical level of texts, but the lexical profiling method has produced results which are comparable to those of the other methods (Lindqvist et al. 2013). Aside from providing data regarding the lexical level of corpora, the lexical profiling method also has the advantage of providing results concerning the presence, i.e., the coverage of specific groups of words in a corpus (e.g., academic words, technical words, proper names, abbreviations, etc.). Another advantage of the method is that the coverage of the word lists obtained can be compared against the reading and listening comprehension thresholds cited earlier, enabling the determination of what vocabulary sizes are required to reach them.

The results produced using the lexical profiling method are quantitative in nature and, on account of this, the method has been somewhat criticized—Crossley et al.

<sup>1</sup> According to the CEFR scale (Council of Europe, 2001).

(2013) argue that the method may result in a certain amount of information loss by reducing the lexical profiles to numbers only. In spite of such criticism, however, it has been widely used (e.g., in Cobb and Horst 1999; Morris and Cobb 2004; Read and Nation 2006; Douglas 2015, etc.).

### 2.3. Word lists

Many word lists have been used in conjunction with the lexical profiling method—though some, however, more often than others. In this brief review, we will mention only the most important ones, with special emphasis on those used in this paper.

For determining the lexical level of corpora, general frequency word lists can be used. Most of these contain about 2,000 words as they were principally created for ELT purposes. The General Service List (the GSL; West 1953) was successfully replaced by Brezina and Gablasova's New GSL (2013) and the New GSL by Browne, Culligan and Philips (2013). These three lists are of a very similar size but, while the old GSL was manually extracted from a 5-million-word corpus, the New GSLs were derived from mega-corpora using computer software. Brezina and Gablasova (2013) combined four corpora (twelve billion words) and applied the criterion of average reduced frequency (calculated from the absolute frequency and the distribution of words). Browne et al. (2013) used a subsection of the Cambridge English Corpus (273 million words) and applied the frequency criterion. Both the new lists were reported by their authors as being more successful than the old GSL in contemporary corpora, which was to be expected given that they were derived from modern corpora of a substantial size.

Given the limited size of these lists, another word list, or rather a set of them, is typically used for determining the lexical level of corpora. It is the set produced by Nation (2006a; 2012; 2017) from a corpus combining the British National Corpus (BNC) and the Corpus of Contemporary American English (COCA) (450 million words altogether). The 2012 version of this set contains twenty-five lists, each comprising 1,000 word families, ordered according to their frequency (the first list represents the most frequent 1,000 words of English, the second contains the second most frequent 1,000 words, etc.). Additional supplementary lists containing proper names, abbreviations, compound nouns without a hyphen and marginal words (letters of the alphabet, swear words, etc.), are also provided.

For profiling specialized vocabulary in a corpus, researchers have many word lists standing at their disposal. A number of them represent either general or specialized academic language. Amongst them, the most famous is the Academic Word List (the AWL; Coxhead 2000), derived from a written academic corpus of 3.5 million words. It contains 570 frequent academic words which are not high-frequency words in general terms (as represented by the GSL). Another general academic word list is the Academic Vocabulary List (the AVL; Gardener and Davies 2014), which was built 'from scratch' from a 120-million-word corpus by applying the keyness criterion (the words occurring

much more often in a specialized rather than a general reference corpus). Both Newman (2016) and Hernandez (2017) determined that the AVL performs better than the old AWL, even though such comparisons are exceedingly difficult to make, given that these word lists were made using different criteria and for different groups of learners (the AWL was built under the presumption that its target users already know the most frequent 2,000 words of English, while the AVL makes no such assumption). Recently, the Academic Spoken Word List (ASWL) was also created (Dang et al. 2017), based on a corpus of 13 million words from twenty-four different fields. This word list has 1,741 word families and features a coverage of about 90% in the corpus of its origin. This word list, like the AVL, was also built ‘from scratch’, i.e., no group of words was excluded.

#### 2.4. Vocabulary of university lectures

Dang and Webb (2014) lexically profiled the British Academic Spoken English Corpus (BASE; Thompson and Nesi 2001), which consists of 160 lectures and 40 seminars. Their results are not only valid for academic lectures, but they should provide a close indication of their lexical profile. Dang and Webb found that 4,000 words with proper names and marginal words were necessary to reach a coverage of 95% (corresponding to ‘good’ listening comprehension), while a vocabulary of 8,000 words was necessary for a coverage of 98% (‘optimal’ listening comprehension) in the BASE corpus. Further, the AWL accounted for 4.41% of this corpus.

As for ELF lectures, to the best of our knowledge, no lexical profiling study of the type described above exists. ELF lectures have, though, been studied using different linguistic methods. Mauranen (2010; 2012) studied the English as a Lingua Franca (ELFA) academic corpus, which includes twenty university lectures, and found that it frequently featured strategies aimed at overcoming potential communication problems. These strategies included both simplification and complexification, as well as a pronounced tendency to cooperation, explicitness, unconventional linguistic expressions and substantial metadiscourse to express both evaluation and vagueness, *inter alia*. Mauranen (2012) found that in ELFA “bilingual processing biases lexical choices toward the most frequent, presumably the most deeply entrenched, vocabulary” and argues that the same happens in translations. A major lexical process was found to be what she calls *approximation* (Mauranen 2012), which reflects what happens when ELF speakers approximate the target language expressions and end up with creative word forms and uses. Ranta (2018, 250) found that, in terms of grammar, “non-standard features in ELF do not, as a rule, cause misunderstandings in communication.” Most studies on ELF describe various phenomena related to its variability and contextual adaptability (Cogo & House 2018; Osimk-Teasdale 2018, etc.), which includes variability in vocabulary and how vocabulary is selected and strategically explained. Our study aims to discover whether these ELF specificities will reflect on the lexical profile of international university lectures.

### 2.5. Teaching English from an ELF perspective

Given that our study aims to produce some pedagogically relevant implications, we will say a few words here about pedagogically oriented ELF-studies. Several authors have proposed that for teaching English with an ELF-perspective or ELF-orientation, teachers should be made aware of the principles arising from ELF research and how these have a bearing on their teaching contexts (Bayyurt and Sifakis 2015; Sifakis 2019). We can, in this sense, also speak about ELT which is ELF-aware (Sifakis 2019) or ELF-informed (Seidlhofer 2015). Unlike focusing on codified and stable English forms in typical normative ELT, ELT with an ELF-perspective should have a focus on communication strategies (CSs) (Cogo and Dewey 2012; Jenkins 2014) and is, in a way, post-normative (Dewey 2012). Based on the empirical findings arising from the analysis of ELF corpora—such as ELFA (2008), VOICE (2011) and ACE (Kirkpatrick 2010)—teaching CSs in ELT with an ELF-perspective has been advocated for, given their extremely important function in the co-construction of meaning (Björkman 2014; Jenkins 2014; Vettorel 2018; Sifakis 2019, etc.).

Some of these CSs have been mentioned in 2.4 above; a detailed overview of CSs studied in ELF contexts is given in Björkman (2014). Due to lack of space, we shall mention just one of their possible classifications—for instance, Vettorel (2018) adapts the CSs framework of Celce-Murcia, Dörnyei and Thurrell (1995) to findings from ELF research and defines four macro-areas of ELF CSs: appeal for help, meaning negotiation (requests for repetition, clarification, confirmation checks, etc.), responses (rephrasing, simplification, confirmation, repair, rejection, lexical anticipation/suggestion/correction, use of fillers and time-gaining devices) and achievement strategies (paraphrase, approximation, restructuring, word-coinage, code-switching, translation, foreignizing, etc.). If these are found to significantly influence the lexical profile of international lectures, this will additionally corroborate the above views which point to the need for focusing on them for ELF-oriented ELT.

### 3. DATA AND METHOD

To represent international lectures in English, two corpora are used in this study:

– The ELFA corpus (2008)—the best-known ELF spoken academic corpus. The entire corpus consists of about one million words of academic spoken interaction from four Finnish universities. There are twenty graduate lectures from various disciplines (behavioral sciences, humanities, medicine, social sciences, technology, economics and administration and natural sciences) in this corpus and they together provide 141,502 tokens.<sup>2</sup> Nineteen lecturers speak English as L2 and for one lecturer English is their mother language (Jamaican English).

---

<sup>2</sup> The word counts referred to here for this and other corpora used are those provided by AntWordProfiler (Anthony 2014); tag annotations are not included in the counts.



– The *Transnational Alignment of English Competences for University Lecturers* (TAEC) corpus (2020), which consists of thirty undergraduate and graduate ELF lectures from five European universities (Croatia, Denmark, Italy, the Netherlands and Spain). The lectures come from three broad disciplines: social sciences and humanities, life and medical sciences and physical sciences and engineering. This corpus comprises 343,347 tokens. All lecturers from the corpus speak English as L2.

By combining the ELFA and the TAEC corpora, we obtained a corpus of fifty international lectures in English (484,849 tokens), representing different disciplines, from nine European universities.

To represent U.S. lectures in English, we use the Michigan Corpus of Academic Spoken English (MICASE) (Simpson et al. 1999), i.e., the lecture subsection of this corpus, which is of a comparable size to the international lectures corpus outlined above. This corpus includes sixty-two undergraduate and graduate lectures recorded at the University of Michigan (617,502 tokens), representing various disciplines: social sciences, biological sciences, humanities and arts and physical sciences and engineering. In only one of the lectures, the lecturer's first language is not English. The context in which MICASE lectures were delivered is one where English as a native language is the norm (Mauranen 2010, 185).

We use two computer programs for this study. The main one is AntWordProfiler 1.4.1 (Anthony 2014). This program is recommended by Nation (2017) as the best currently available program for vocabulary profiling, given that the Range program, which was widely used previously, has not been updated for many years. The second program is called AntConc 3.5.8 (Anthony 2019), which is here used to determine keywords—words used much more frequently in the corpus under study than in a given reference corpus. The default settings in this program for determining keywords were applied—keyword statistic: log-likelihood (4-term), keyword statistic threshold ( $p < 0.05$  (+Bonferroni)), keyword effect size measure (dice coefficient).

The coverages representing the good and the ideal listening comprehension levels are those quoted earlier, i.e., those reported by Van-Zeeland and Schmitt (2012): 95% and 98% respectively, which were supported by Laufer and Ravenhorst-Kalovski (2010), and also used by Dang and Webb (2014) in their profiling of the BASE corpus.

The word lists used for profiling general vocabulary will be Nation's (2012). His word list set is the most comprehensive and in fact the only one which can be used for mid- and low-frequency words, given that other general word lists include only high-frequency words. To represent academic vocabulary, the obvious choice is the ASWL (Dang et al. 2017), as it is a recent, carefully generated list, which represents spoken academic English. However, as this word list captures only high-frequency vocabulary encountered in academic speech and is thus very general in nature, we additionally use the AWL by Coxhead (2000), which represents mid-frequency academic vocabulary, i.e., academic words beyond the most frequent 2,000 words of English, based on English as a Native Language (ENL) academic texts. The AWL

has been widely used for two decades now and the results for its coverages in various corpora are available for comparison (Chen and Ge 2007; Vongpumivitch et al. 2009; Valipouri and Nassaji 2013, etc.).

#### 4. RESULTS AND ANALYSIS

This section is divided into three parts, each answering one of the research questions.

4.1. Given ELF specificities, is the general lexical profile of international academic lectures different from that of U.S. academic lectures in English?

Table 1 provides the coverages for Nation's word list set in the corpus of international lectures (the ELFA and the TAEC), as well as in the reference corpus of the U.S. lectures (MICASE). The coverages related to the threshold for good listening comprehension (95%) are given in bold, while those relating to ideal listening comprehension (98%) are underlined. The same as Dang and Webb (2014), we assume that proper names and marginal words (the latter chiefly includes the letters of the alphabet) carry no comprehension burden for the listener; we additionally assume that abbreviations also do not carry additional comprehension load for the listener, as these are chiefly explained the first time they are used and may be understood from the context. Tag annotations and strings of numbers and letters were discarded from the calculations.

TABLE 1. Lexical Frequency Profile of International and U.S. Lectures in English (%)

BNC/COCA word lists	ELFA+TAEC lectures (cum. coverage %)	MICASE lectures (cum. coverage %)
Proper names	1.24	1.43
Marginal words	5.4	4.58
Abbreviations	5.54	4.81
1,000	83.52	86.01
2,000	90.04	91.71
3,000	94.31	<b>95.38</b>
4,000	<b>95.45</b>	96.75
5,000	96.12	97.38
6,000	96.58	97.82

BNC/COCA word lists	ELFA+TAEC lectures (cum. coverage %)	MICASE lectures (cum. coverage %)
7,000	96.87	<u>98.16</u>
8,000	96.99	98.37
9,000	97.12	98.51
10,000	97.23	98.69
11,000	97.35	98.82
12,000	97.45	98.89
13,000	97.55	98.97
14,000	97.62	99.04
15,000	97.67	99.09
16,000	97.74	99.15
17,000	97.79	99.19
18,000	97.82	99.21
19,000	97.85	99.23
20,000	97.88	99.25
21,000	97.9	99.26
22,000	97.92	99.28
23,000	97.93	99.3
24,000	97.93	99.3
25,000	97.96	99.31
Transparent compounds	<u>98.21</u>	99.4
Off-the-lists items	1.81	0.6

As can be seen in Table 1, high-frequency general vocabulary, which in the literature is taken to be the most frequent 2,000 or 3,000 words of English, covers less vocabulary used in the international than in the U.S. lecture corpus. Namely, proper names, marginal words, abbreviations and the first 2,000 words of English cover 90.04% in the international lectures and 91.71% in the U.S. lectures, while these three supplementary lists and the first 3,000 words of English cover 94.31% of the international lectures as compared to 95.38% of the U.S. lectures. The last figure reaches the threshold for good listening comprehension in the MICASE corpus. Differently from this, Dang and

Webb (2014) concluded that 4,000 words were needed to listen to lectures and seminars from the BASE corpus, but our calculations for just the BASE lectures confirmed the figure we obtained for the MICASE lectures, i.e., 3,000 words for good listening comprehension. Our finding is also in line with Nation's (2006) result that 3,000 words (plus proper names and marginal words) are typically required to understand general spoken English.

The level of good listening comprehension in the international lectures is reached at 4,000 most frequent words of English, together with proper names, marginal words and abbreviations (95.45%). This is 1,000 words more than is needed for the U.S. lectures. The level of 3,000 words, needed for the U.S. lectures, is the level of B1 students (according to the CEFR, and as mentioned in Section 2.1), while B2 level is required for comprehension of the international lectures.

The level of ideal listening comprehension is reached at 7,000 most frequent words of English for the MICASE lectures, somewhat fewer than found by Dang and Webb (2014) for British academic speech in general, which was 8,000 words. Our figure is again in line with Nation's finding for general spoken English, which he estimated to require some 6,000-7,000 words for ideal-level comprehension. A vocabulary size of 7,000 words is certainly the level of non-native students who are proficient in English, such as non-native postgraduates pursuing doctoral studies in English (Nation, 2013).

The coverage threshold needed for the ideal level of listening comprehension could only be reached using all Nation's lists in the international lectures corpus—all Nation's lists exhaust 98.21% of the words in them, compared to 99.4% of the words in the MICASE lecture corpus. By inspecting the off-the-lists items in the international lectures corpus, we determined that many were proper names which Nation's names list did not capture as it had been based on the U.K. and the U.S. corpora (e.g., *Bakharev, Bjornenak, Wald, Tsyganenko, Trogir, Tampere...*), words from languages other than English (e.g. *zvijezda, ragazzi, sorta, cedevita, exemple, authorita, certificada, comunque, quindi, scusate, credi...*), approximations (Mauranen 2012) such as creatively inflected forms of some words, not encompassed by the word lists (e.g., *informations, digged, sticked, binded, bringed, detailds, deviced*) or creative derivatives, also not encompassed by the word lists (e.g., *territoral, degradate, explainment, organizatory, unconsistent, unvoluntary*). This vocabulary is reflective of the use of CSs such as translation, code-switching, creative word-coinage and foreignizing, which have already been widely documented in the ELF literature (Pitzl et al. 2008; Breiteneder 2009; Mauranen 2018; Ranta 2018, etc.). It must be pointed out that in any corpus there will always be some items which are off the lists, for several reasons: first, no word list set can ever be exhaustive—e.g., Nation (2013) estimates that there are as many as 70,000 word families in English, while his lists represent only the most frequent 25,000. Second, there will always be some new proper names and newly coined words. Finally, there will always be some misspellings, typos, errors and

creative new forms in any corpus. All corpora used in this study have been carefully produced and reviewed, and so the number of errors should be reduced to a minimum and therefore not affect the lexical profiling results.

We should not automatically assume that the words exemplified in the paragraph above will not be understood by students listening to an ELF lecture; on the contrary, we can expect that in many cases they will understand them given the shared context of the lesson and perhaps the shared mother tongue of lecturer and students. For instance, Ranta (2018, 251) finds that misunderstandings on account of grammatical form in ELF are quite rare. Indeed, many of the creative word forms found in such contexts are “semantically transparent” (Björkman 2018) while Hynninen and Solin (2018, 267) argue that “[n]on-conformity to ENL may be functionally motivated and even enhance mutual understanding.” Bearing this in mind, it can reasonably be expected that in international lecture situations the level of good and ideal listening comprehension will be lower than those detailed above, although it is difficult to pinpoint exactly the vocabulary size needed to follow an international lecture in English. We attempted to make this estimate by inspecting the off-the-lists items. The international lectures corpus contained 8,223 such items; instead of analysing the whole list, which contains words from six different languages, we carefully inspected four 300-word subsets. These 1,200 off-the-lists items covered 0.67% of the complete corpus. Based on the assumption that these will be understandable to international students, we added said coverage to the figures in Table 1 and arrived at the following conclusions:

- a) The level of good listening comprehension is reached at 3,000 words in international lectures in English, which is the same as in U.S. lectures.
- b) 11,000 words are needed for the level of ideal comprehension in international lectures in English, which is substantially different from U.S. lectures.

The lexical profile of international lectures in English, therefore, displays great similarities to that of U.S. lectures in English, in terms of how often they use the most frequent English vocabulary, but there is also much greater variation in the words used beyond the basic core vocabulary.

#### 4.2. What are the differences in the use of high- and mid-frequency words between international and U.S. academic lectures in English?

To further compare the lexical profiles of international and U.S. lectures in English, we determined which words were used far more frequently in the international than in the U.S. lectures; then we classified them according to the word frequency band they belong to. Table 2 shows the high- and mid-frequency words which are most frequent in the international lectures corpus, low-frequency words (those beyond the most frequent 9,000 words of English) generally being technical and more likely to reflect the specific topics discussed in the corpus.

TABLE 2. High- and Mid-Frequency Words Used Significantly more Frequently in International than in U.S. Lectures in English

BNC/COCA 1 <sup>st</sup> 1,000 words							
be	part	bottle	draw	high	noise	record	telephone
company	present	bright	during	hour	normal	relate	television
market	price	business	easy	important	north	responsible	then
tax	programme	but	educate	inform	not	rights	this
yes	read	buy	explain	inside	on	ring	throat
bank	return	cause	field	insure	or	see	total
can	service	centre	for	internet	other	shoot	touch
cost	situation	change	from	know	pain	should	understand
country	student	clear	go	laugh	perfect	side	very
employ	we	colour	green	manage	please	smoke	weight
govern	already	couple	group	many	point	so	which
learn	also	course	hand	maybe	possible	sorry	why
main	because	difference	have	minute	post	strong	will
name	bit	different	hear	need	quite	system	wine
nation	blood	difficult	heart	neighbour	radio	tail	zero
ok	board	doctor	here	no	reason		
BNC/COCA 2 <sup>nd</sup> 1,000 words							
organize	model	calculate	disease	interview	nail	release	spin
finance	nerve	cap	economy	justice	nervous	research	split
library	plane	cent	electric	labour	non	risk	staff
centimetre	product	chain	example	male	observe	role	standard
contribute	slide	chat	fantastic	mathematics	patient	rope	super
cough	stock	condition	favour	medical	per	screen	tour
customer	vary	connect	flight	medicine	perform	seconds	trunk
emotion	account	contain	foreign	mirror	pole	silence	twin
fruit	active	county	hedge	modern	process	skill	union
fund	bin	current	hotel	motor	propose	society	university
future	bond	develop	instance	mouse	quality	soil	wage
improve	brain	discuss	instrument	muscle	receive	specific	wipe
metre	brand						
BNC/COCA 3 <sup>rd</sup> 1,000 words							
budget	protein	climate	effective	gender	link	organic	relative
institution	revenue	colleague	enterprise	highway	magnet	participant	satisfaction
aim	structure	component	etc	implement	manipulate	patients	sector
cluster	academy	compromise	evaluate	initiate	margin	philosophy	solve
complex	airline	confer	excess	internal	mild	principle	summit
correlate	analyse	constitution	extract	international	molecule	quantity	supervise
gesture	behaviour	crystal	fibre	invest	nod	ratio	sympathy
innovate	border	dimension	formula	journalist	obtain	rational	transport
objective	candidate	division	frequent	liberal	organ	recruit	yield
principle	capacity	DNA	gap				

BNC/COCA 4 <sup>th</sup> 1,000 words							
click	activate	cord	electron	impair	monopoly	plague	spine
orbit	axe	diabetes	frontier	infrastructure	municipal	rotate	synthesis
portfolio	civic	distort	graph	laser	originate	sodium	thesis
residue	configure	duration	hydrogen	leaf	periphery	sparkle	tumour
BNC/COCA 5 <sup>th</sup> 1,000 words							
symmetry	bail	condense	gram	membrane	par	pragmatic	sulphur
applause	beta	deviate	id	nowadays	plasma	sniff	temporal
axis	cellular	diversify	kidney				
BNC/COCA 6 <sup>th</sup> 1,000 words							
euro	calcium	duct	inclusive	notch	respirator	scroll	sic
asymmetry	cleave	dummy	intern	oscillate	scissors	sensory	soda
bladder	degenerate	excel	neuron				
BNC/COCA 7 <sup>th</sup> 1,000 words							
antenna	cohesion	harmonic	null	potassium	que	RNA	teller
auxiliary	electrode	methane	placebo	pulp	resin		
BNC/COCA 8 <sup>th</sup> 1,000 words							
peptide	amplitude	autistic	commune	stagnate	surname		
BNC/COCA 9 <sup>th</sup> 1,000 words							
acupuncture	crystalline	kilo	mitochondria	vertebra	visceral		

Table 2 contains 398 words, which cover 30.65% of the international lecture corpus and just 20.76% of the MICASE lecture corpus. A difference of 10% for just 398 words is indeed substantial and points to a difference in the vocabulary profiles of the two corpora, which has its pedagogical implications. Only a minority of the words above are included in the table simply on account of them reflecting the particular topics discussed in the international lectures corpus, most of them in fact seeming to be quite general; the list contains important function words, academic words and formal vocabulary, which were used far more frequently in the international than in the U.S. lectures. Given its brevity, the word list from Table 2 could itself be used as a resource in teaching English with an ELF-perspective (after eliminating some of the words specific to the lecture topics, as argued above).

The 1,000 most frequent words of English include a number of function words, such as auxiliaries (*be, can, have, will, should*), some basic conjunctions (*but, because, so, or*), pronouns (*we, which*), prepositions (*on, for, from, during*), quantifiers (*many*), pro-sentences and negations (*yes, ok, no, not*) and demonstratives (*this*), and these were used much more often in the international than in the U.S. lectures. On the one hand, ELF language is prone to simplification, which is itself an ELF CS (Mauranen 2016; 2018) and is often reflected in a high concentration of very frequent words, including function words. This phenomenon is also noted in translation studies, which is why Mauranen (2016, 36) suggests that it may be a general contact-induced phenomenon. On the other hand, ELF also tends towards complexification (Mauranen 2018), another ELF CS, which is here reflected in the greater

overall variation with respect to low-frequency words, as seen above.

The following fifty-eight words from Table 2 are part of the Academic Word List (Coxhead 2000): *academy, activate, analyse, bond, capacity, colleague, complex, component, confer, constitution, contribute, couple, deviate, dimension, distort, diversify, duration, economy, evaluate, extract, finance, formula, fund, gender, implement, infrastructure, initiate, innovate, instance, institution, internal, invest, labour, liberal, link, manipulate, margin, medical, normal, objective, obtain, participant, philosophy, principal, principle, process, ratio, rational, release, research, revenue, role, sector, specific, structure, thesis, transport, vary*. These were particularly frequent in our international lectures corpus and we consider that perhaps pedagogical focus should be given to them in teaching academic English with an ELF-perspective.

Some items from Table 2 point to an overabundance of certain deictic words (e.g., *then, here, this, go*). Proximal deictic expressions such as *this* and *here* (also the pronoun *we*) frequently accompany pointing gestures, which are frequent in contact phenomena and once again represent an ELF CS. Mauranen (2018) also found that expressing evaluations was frequent in academic ELF, which may explain why words such as *quite, very, perfect, fantastic, big, mild* and *evaluate* were used far more frequently in the international lectures corpus compared to MICASE.

We now turn our attention to how frequent academic vocabulary is used in the international and the U.S. lectures corpora.

#### 4.3. How much high- and mid-frequency academic vocabulary is present in international compared to U.S. academic lectures in English?

We first provide the coverages reached by the ASWL (Dang et al. 2017) in our two corpora (Table 3). As explained before, this list was built ‘from scratch’ and contains high-frequency words commonly found in academic speech.

TABLE 3. Coverage of High-Frequency Academic Vocabulary in International Compared to U.S. Lectures in English (%)

Word list	ELFA+TAEC lectures (coverage %)	MICASE lectures (coverage %)
ASWL	84.4	86.95

The ASWL achieves a somewhat lower coverage in the international lecture corpus than in the MICASE lecture corpus. This indicates that the ASWL is a useful resource for academic ELF, but not as useful as it is for U.S. lectures.

We will also examine the coverage of the AWL (Coxhead 2000) as, although this academic list is based on a written corpus, it may offer an insight into the distribution



of the more specific academic vocabulary in the two corpora. As said earlier, the AWL contains mid-frequency words since it was built on the basis that the 2,000 most high-frequency words (represented by the GSL) were excluded. As it was built on the back of the GSL, the coverages of the two are typically shown together (Table 4).

TABLE 4. Coverage of Mid-Frequency Academic Vocabulary in International Compared to U.S. Lectures in English (%)

Word list	ELFA+TAEC lectures (coverage %)	MICASE lectures (coverage %)
GSL	80.64	80.4
AWL	4.19	3.76
GSL+AWL	84.83	84.16

While the international lectures corpus contains fewer high-frequency academic words than the MICASE lecture corpus (Table 3), the situation with the mid-frequency academic vocabulary is reversed (Table 4)—the AWL coverage is somewhat greater in the international lectures than in the U.S. lectures. The result obtained for the international lectures is in fact quite close to what Dang and Webb (2017) found for British academic speech, which was 4.41%. We again point to the finding that some words from the AWL were disproportionately used more frequently in the international than in the U.S. lectures, a finding that clearly has pedagogical implications.

## 5. DISCUSSION AND PEDAGOGICAL IMPLICATIONS

Our first research question was how the general lexical profile of the international lectures in English differed from that of the U.S. lectures in English, given the specificities of ELF. We determined that the two profiles were similar in some respects, but very different in the others. The level of good listening comprehension was calculated to be 3,000 words (plus proper names, marginal words and abbreviations) for the U.S. lectures, while 4,000 words were needed for the international lectures. However, we took into account that the proper names list did not capture many names from the international lectures, as well as the fact that many words were missed by the list because they had creatively derived or inflected forms. These variations, as suggested earlier, do not hinder comprehension. In addition, words from languages other than English were present in the international lectures, indicative of the use

of the ELF CSs of translation, code-switching and foreignization, and one can only assume that the lecturer used them expecting them to be understood by the students. On the basis of this assumption, the calculations yielded the same figure for the international as for the U.S. lectures—3,000 words were needed to follow them at the good level.

On the one hand, some 7,000 of the most frequent words of English (plus proper names, marginal words and abbreviations) were needed to reach ideal comprehension in the U.S. lectures. On the other hand, without the assumption made above, the level of ideal comprehension was scarcely achieved in the international lectures corpus even with all the word lists available. However, under said assumption, the level of ideal comprehension could be reached at 11,000 most frequent words of English in the international lectures corpus. This suggests differences in mid- and low-frequency words in the international lectures compared to the U.S. lectures. The existence of these differences reflects the extensive use of certain ELF CSs, as pointed out above, and corroborates the need to focus on these in teaching academic English with an ELF-perspective, as has already been suggested in the literature (Cogo and Dewey 2012; Björkman 2014; Jenkins 2014; Bayyurt and Sifakis 2015; Vettorel 2018; Sifakis 2019, etc.). Additionally, the results may also point to the need to create special word lists for ELF lectures, in order to provide a greater coverage, i.e., the percentage of the words that a student understands, with fewer words. As things stand now, a learner has to master 11,000 words to follow ELF lectures at the ideal level, although this could presumably be done with fewer words selected based on their frequency in the ELF corpus. However, the value of a word list which would contain such words, given the fluidity of ELF (Seidlhofer 2011) depending on the context in which it takes place, would need to be investigated.

In response to the second research question, we drafted a word list containing 398 high- and mid-frequency words which were disproportionately used more often in the international than in the U.S. lectures in English. The results again reflected the extensive use of certain ELF CSs. It was demonstrated that some simple function words were much more concentrated in the international lectures than in the U.S. ones. Also, naturally, some of the words on the list reflected the particular topics discussed in the lectures. The list could be used as a resource for teaching academic English with an ELF-perspective (after the elimination of lecture topic-specific words), as it ensures 10% greater coverage in ELF lectures than in U.S. lectures. However, as suggested above, its value would have to be examined in different ELF contexts.

In response to the third research question, we found that the international lectures in English covered fewer high-frequency academic words than the U.S. lectures, but more specific, mid-frequency academic words. About 10% of the AWL was found to be disproportionately more used in the international than in the U.S. lectures. This list of about sixty academic words should also be given special attention when teaching academic English with an ELF-perspective.

## 6. LIMITATIONS OF THE STUDY

This study is solely focused on the vocabulary required for listening comprehension, given that vocabulary knowledge is an important predictor of comprehension. Given its limited scope, it does not take into account other factors which affect listening comprehension.

By combining the ELFA and the TAEC corpora, we obtained a corpus of a reasonable size for lexical profiling, although the tendency is to do this type of study on much larger corpora. That said, we are of the opinion that it is important to do a profiling study on ELF lectures with the available corpora (which are, in effect, sufficiently large), rather than wait for more ELF lecture corpora to become available in the future, as there are important pedagogical implications that arise from the results of such studies.

One U.S. academic context was chosen as representative for Anglophone contexts for the purpose of this study. We believe that choosing another Anglophone context, for instance British, would not largely affect the general findings of the study with regard to the vocabulary differences found which arise from the specific use of CSs in ELF contexts. Namely, given that ELF is not a variety of English but rather a mode of English (Jenkins 2015, 54-55; Mauranen 2018, 9; Bulatović 2021), it is not modelled on any native-speaker standard, i.e., it is not closer to any particular English variety (for instance, British or American) and farther from others. Instead, ELF is rather endonormative (Schmitz 2012), governed by its own intrinsic rules and specificities, which are reflected in its lexical profile. The size of MICASE makes that corpus comparable to the available ELF corpora and for this reason, as well as those above, many comparative studies of ELF have no problem in using this particular corpus as their sole reference corpus—for instance, in much of Mauranen's work (2012; 2018, etc.).

## 7. CONCLUSION

In response to the first research question (*Given the ELF specificities, is the general lexical profile of international academic lectures different from that of U.S. academic lectures in English?*), we showed that 3,000 words are needed for good listening comprehension in terms of both international and U.S. lectures in English, while ideal comprehension is achieved at 11,000 words for the former and 7,000 words for the latter, suggesting differences between the two, mostly with respect to low-frequency vocabulary variation. Many of the differences came from the extensive use of ELF CSs in international lectures, such as translation, code-switching and foreignization.

In response to the second research question (*What are the differences in the use of high- and mid-frequency words between international and U.S. academic lectures in English?*), we drafted a word list containing some 400 high- and mid-frequency words which were disproportionately used more often in the international lectures than in the U.S. lectures. The results again reflected the extensive use of particular ELF CSs, such as simplification, complexification and use of gestures.

In responding to the final research question (*How much high- and mid-frequency academic vocabulary is present in international compared to U.S. academic lectures in English?*), we found that international lectures also featured less high-frequency and more mid-frequency academic vocabulary than U.S. lectures. It was also established that international lectures in English abound in academic vocabulary and that the existing academic word lists are useful resources for those intending to listen to such lectures. However, vocabulary word lists that are specific to academic ELF might be even more useful, as the distribution of academic words in international lectures in English differs somewhat from that in U.S. lectures; the value of such resources would, however, have to be corroborated in different ELF contexts.

The results of this study show some differences in the lexical profile of international and U.S. lectures in English which have important pedagogical implications. While, on the one hand, they highlight the importance of focusing on ELF CSs in ELT with an ELF-perspective, on the other they also point to the conclusion that it might be worthwhile to draft vocabulary lists which are more specific to ELF.

#### WORKS CITED

- ANDERSEN, Richard C. and Peter Freebody. 1981. "Vocabulary Knowledge." In Guthrie 1981, 77-117.
- ANTHONY, Laurence. 2014. *AntWordProfiler (Version 1.4.1)* [computer software]. Tokyo: Waseda University.
- . 2019. *AntConc (Version 3.5.8)* [computer software]. Tokyo: Waseda University.
- BARDEL, Camilla, Cristina Lindqvist and Batia Laufer, eds. 2013. *L2 Vocabulary Acquisition, Knowledge and Use: New Perspectives on Assessment and Corpus Analysis*. EuroSLA.
- BARTNING, Inge, Maisa Martin and Ineke Vedder, eds. 2010. *Communicative Proficiency and Linguistic Development: Intersections between SLA and Language Testing Research*. Eurosla.
- BAYYURT, Yasemin and Nicos C. Sifakis. 2015. "ELF-aware In-service Teacher Education: A Transformative Perspective." In Bowles and Cogo, 2015, 117-35.
- BJÖRKMAN, Beyza. 2011. "English as a Lingua Franca in Higher Education: Implications for EAP." *Ibérica: Revista de la Asociación Europea de Lenguas para Fines Específicos* 22: 79-100.
- . 2014. "An Analysis of Polyadic English as a Lingua Franca (ELF) Speech: A Communicative Strategies Framework". *Journal of Pragmatics* 66: 122-38.
- . 2018. "Morphosyntactic Variation in Spoken English as a Lingua Franca Interactions: Revisiting Linguistic Variety." In Jenkins et al. 2018, 255-66.
- BOWLES, Hugo and Alessia Cogo, eds. 2015. *International Perspectives on English as a Lingua Franca: Pedagogical Insights*. New York: Palgrave Macmillan.
- BREITENEDER, Angelika. 2009. "English as a Lingua Franca in Europe: an Empirical Perspective." *World Englishes* 28 (2): 256-69.
- BREZINA, Vaclav and Dana Gablasova. 2013. "Is there a Core General Vocabulary? Introducing the New General Service List." *Applied Linguistics* 36 (1): 1-22.

- BROWNE, Charles, Brent Culligan and Joseph Phillips. 2013. "The New General Service List." <http://www.newgeneralservicelist.org/> [Accessed September 1, 2019].
- BULATOVIĆ, Stefan. 2021. *Applying the Concept of English as a Lingua Franca in a Non-Anglophone Academic Context: The Case of Montenegrin Higher Education: Preliminary Research*. Nikšić: University of Montenegro.
- CAPEL, Annette. 2012. "Completing the English Vocabulary Profile: C1 and C2 Vocabulary." *English Profile Journal* 3 (1): 1-14.
- CELCE-MURCIA, Marianne, Zoltan Dörnyei and Sarah Thurrell. 1995. "Communicative Competence: A Pedagogically Motivated Model with Content Specifications". *Issues in Applied Linguistics* 6 (2): 5-35.
- CHEN, Qi and Guang-chun Ge. 2007. "A Corpus-Based Lexical Study on Frequency and Distribution of Coxhead's AWL Word Families in Medical Research Articles (RAs)." *English for Specific Purposes* 26 (4): 502-514.
- COBB, Tom and Marlise Horst. 1999. "Vocabulary Sizes of Some City University Students." *Journal of the Division of Language Studies of City University of Hong Kong* 1: 59-68.
- COGO, Alessia and Martin Dewey. 2012. *Analysing English as a Lingua Franca: A Corpus-Driven Investigation*. New York: Continuum.
- and Juliane House. 2018. "The Pragmatics of ELF." In Jenkins et al. 2018, 210-23.
- COUNCIL OF EUROPE. 2001. *The Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. Cambridge: Cambridge University Press.
- COXHEAD, Averil. 2000. "A New Academic Word List." *TESOL* 34 (2): 213-38.
- , Thi Ngoc Yen Dang and Shota Mukai. 2017. "Single and Multi-Word Unit Vocabulary in University Tutorials and Laboratories: Evidence from Corpora and Textbooks." *Journal of English for Academic Purposes* 30: 66-78.
- CROSSLEY, Scott A., Tom Cobb and Danielle S. McNamara. 2013. "Comparing Count-Based and Band-Based Indices of Word Frequency: Implications for Active Vocabulary Research and Pedagogical Applications." *System* 41 (4): 965-81.
- DANG, Thi Ngoc Yen. 2018. "The Nature of Vocabulary in Academic Speech of Hard- and Soft-Sciences." *English for Specific Purposes* 51: 69-83.
- , Averil Coxhead and Stuart Webb. 2017. "The Academic Spoken Word List." *Language Learning* 67 (4): 959-97.
- and Stuart Webb. 2014. "The Lexical Profile of Academic Spoken English." *English for Specific Purposes* 33 (1): 66-76.
- DEWEY, Martin. 2012. "Towards a Post-Normative Approach: Learning the Pedagogy of ELF". *Journal of English as a Lingua Franca* 1 (1): 141-70.
- and Inmaculada Pineda. 2020. "ELF and Teacher Education: Attitudes and Beliefs". *ELT Journal* 74 (4): 428-41.
- DOUGLAS, Scott Roy. 2015. "The Relationship between Lexical Frequency Profiling Measures and Rater Judgements of Spoken and Written General English Language Proficiency on the CELPIP-General Test." *TESL Canada Journal* 32 (9): 43-64.

- DUFF, Dawna and Melissa Brydon. 2020. "Estimates of Individual Differences in Vocabulary Size in English: How many Words are Needed to 'Close the Vocabulary Gap'?" *Journal of Research in Reading* 43 (4): 454-81.
- ELFA. 2008. *The Corpus of English as a Lingua Franca in Academic Settings*. Director: Anna Mauranen. <http://www.helsinki.fi/elfa> [Accessed August 1, 2021].
- GALLOWAY, Nicola. 2018. "ELF and ELT Teaching Materials." In Jenkins et al. 2018, 468-80.
- GARDENER, Dee and Mark Davies. 2014. "A New Academic Vocabulary List." *Applied Linguistics* 35 (3): 305-27.
- GUTHRIE, John T., ed. 1981. *Comprehension and Teaching: Research Reviews*. Newark, DE: International Reading Association.
- HERNANDEZ, Michelle Morgan. 2017. *Comparing the AWL and AVL in Textbooks from an Intensive English Program*. Theses and Dissertations. 6458. <https://scholarsarchive.byu.edu/etd/6458>
- HINKEL, Eli, ed. 2017. *Handbook of Research in Second Language Teaching and Learning, Vol. III*. New York: Routledge.
- HYNNINEN, Niina and Anna Solin. 2018. "Language Norms in ELF." In Jenkins et al. 2018, 267-78.
- JENKINS, Jennifer. 2014. *English as a Lingua Franca in the International University*. New York: Routledge.
- . 2015. "Repositioning English and Multilingualism in English as a Lingua Franca". *English in Practice* 2 (3): 49-85.
- , Will Baker and Martin Dewey, eds. 2018. *The Routledge Handbook of English as a Lingua Franca*. New York: Routledge.
- KIRKPATRICK, Andy. 2010. "Researching English as a Lingua Franca in Asia: The Asian Corpus of English (ACE) Project". *Asian Englishes* 13 (1): 4-18.
- LAUFER, Batia. 1989. "What Percentage of Text Lexis is Essential for Comprehension?" In Laurén and Nordman 1989, 316-23.
- and Paul Nation. 1995. "Vocabulary Size and Use: Lexical Richness in L2 Written Production." *Applied Linguistics* 16 (3): 307-22.
- and Geke C. Ravenhorst-Kalovski. 2010. "Lexical Threshold Revisited: Lexical Text Coverage, Learners' Vocabulary Size and Reading Comprehension." *Reading in a Foreign Language* 22 (1): 15-30.
- LAURÉN, Christer and Marianne Nordman, eds. 1989. *Special language: From Humans Thinking to Thinking Machines*. Clevedon: Multilingual Matters.
- LINDQVIST, Christina, Anna Gudmundson and Camilla Bardel. 2013. "A New Approach to Measuring Lexical Sophistication in L2 Oral Production." In Bardel et al. 2013, 109-126.
- MAURANEN, Anna. 2010. "Features of English as a Lingua Franca in Academia." *Helsinki English Studies* 6 (6): 6-28.
- . 2012. *Exploring ELF: Academic English Shaped by Non-Native Speakers*. Cambridge: CUP.

- . 2016. “English as a Global Lingua Franca: Changing Language in Changing Global Academia.” In Murata 2016, 41-58.
- MAURANEN, Anna. 2018. “Conceptualising ELF.” In Jenkins et al. 2018, 7-24.
- MILTON, James. 2010. “The development of Vocabulary Breadth across the CEFR Levels.” In Bartning et al. 2010, 211-32.
- MORRIS, Lori and Tom Cobb. 2004. “Vocabulary Profiles as Predictors of the Academic Performance of Teaching English as a Second Language Trainees.” *System* 32 (1): 75-87.
- MURATA, Kumiko, ed. 2016. *Exploring ELF in Japanese Academic and Business Contexts*. London: Routledge.
- NAGY, William E. 1988. *Teaching Vocabulary to Improve Reading Comprehension*. Newark, DE: International Reading Association.
- NATION, Paul. 2006. *The BNC/COCA Word Family Lists*. <https://www.wgtn.ac.nz/lals/resources/paul-nations-resources/vocabulary-lists>. [Accessed April 1, 2018].
- . 2006b. “How Large a Vocabulary is Needed for Reading and Listening?” *Canadian Modern Language Review* 63 (1): 59-82.
- . 2012. *The BNC/COCA Word Family Lists*. <http://www.victoria.ac.nz/lals/about/staff/paul-nation>. [Accessed April 1, 2018].
- . 2013. *Learning Vocabulary in Another Language (2<sup>nd</sup> Edition)*. Cambridge: Cambridge University Press.
- . 2017. *The BNC/COCA Word Family Lists*. [https://www.wgtn.ac.nz/\\_data/assets/pdf\\_file/0005/1857641/about-bnc-coca-vocabulary-list.pdf](https://www.wgtn.ac.nz/_data/assets/pdf_file/0005/1857641/about-bnc-coca-vocabulary-list.pdf). [Accessed 1 August 2021].
- and Laurence Anthony. 2017. “Measuring Vocabulary Size”. In Hinkel 2017, 355-68.
- NEWMAN, Jacob Andrew. 2016. *A Corpus-Based Comparison of the Academic Word List and the Academic Vocabulary List*. Theses and Dissertations. 6080. <https://scholarsarchive.byu.edu/etd/6080>
- OSIMK-TEASDALE, Ruth. 2018. “Analysing ELF Variability.” In Jenkins et al. 2018, 201-209.
- PITZL, Marie-Luise, Angelika Breiteneder and Theresa Klimpfinger. 2008. “A World of Words: Processes of Lexical Innovation in VOICE.” *Vienna English Working Papers* 17 (2): 21-46.
- RANTA, Elina. 2018. “Grammar in ELF.” In Jenkins et al. 2018, 244-54.
- READ, John and Paul Nation. 2006. “An Investigation of the Lexical Dimension of the IELTS Speaking Test.” *IELTS Research Reports* 6: 207-31.
- SCHMITT, Norbert, Xiangying Jiang and William Grabe. 2011. “The Percentage of Words Known in a Text and Reading Comprehension.” *The Modern Language Journal* 95 (1): 26-43.
- SCHMITZ, John Robert. 2012. “To ELF or not to ELF?(English as a Lingua Franca): That’s the Question for Applied Linguistics in a Globalized World”. *Revista brasileira de linguística aplicada*, 12: 249-84.

- SEIDLHOFER, Barbara. 2011. *Understanding English as a Lingua Franca*. Oxford: OUP.
- . 2015. “ELF-Informed Pedagogy: From Code-Fixation towards Communicative Awareness”. In Vettorel, 2015, 19-30.
- SIFAKIS, Nicos. 2019. “ELF Awareness in English Language Teaching: Principles and Processes.” *Applied linguistics*, 40 (2): 288-306.
- SIMPSON, Rita C., Sarah L. Briggs, Janine Ovens and John Swales. 1999. *The Michigan Corpus of Academic Spoken English*. Ann Arbor, MI: The Regents of the University of Michigan.
- TAEC. 2020. *The Transnational Alignment of English Language Competences for University Lecturers Corpus*. TAEC Erasmus+ project (2017-2020). Director: Slobodanka Dimova.
- THOMPSON, Paul and Hilary Nesi. 2001. “The British Academic Spoken English (BASE) Corpus Project.” *Language Teaching Research* 5 (3): 263-64.
- VALIPOURI, Leila and Hossein Nassaji. 2013. “A Corpus-Based Study of Academic Vocabulary in Chemistry Research Articles.” *Journal of English for Academic Purposes* 12 (4): 248-263.
- VAN ZEELAND, Hilde and Norbert Schmitt. 2013. “Lexical Coverage in L1 and L2 Listening Comprehension: The Same or Different from Reading Comprehension?” *Applied Linguistics* 34 (4): 457-79.
- VETTOREL, Paola, ed. 2015. *New Frontiers in Teaching and Learning English*. Cambridge: Cambridge Scholars.
- . 2018. “ELF and Communication Strategies: Are they Taken into Account in ELT Materials?” *RELC Journal* 49 (1): 58-73.
- VOICE. 2011. The Vienna-Oxford International Corpus of English (Version 1.0 XML). Director: Barbara Seidlhofer; Researchers: Angelika Breiteneder, Theresa Klimpfner, Stefan Majewski, Marie-Luise Pitzl.
- VONGPUMIVITCH, Viphavee, Ju-yu Huang and Yu-Chia Chang. 2009. “Frequency Analysis of the Words in the Academic Word List (AWL) and Non-AWL Content Words in Applied Linguistics Research Papers.” *English for Specific Purposes* 28 (1): 33-41.
- WEST, Michael. 1953. *A General Service List of English Words*. London: Longman.

Received 12 September 2021

Revised version accepted 28 October 2022

Milica Vuković-Stamatović, PhD, is Associate Professor of English Language and Linguistics at the Faculty of Philology, University of Montenegro. She currently teaches Discourse Analysis and English for Specific Purposes—she has taught English for Physics and English for Mathematics for almost two decades. Her main research interests lie in the field of discourse analysis, pragmatics and ESP. She has published in journals such as *Corpus Linguistics and Linguistic Theory*, *Discourse & Society*, *Pragmatics & Society*, *Journal of Language and Politics*, *Gender and Language*, *Lexikos*, among others.