ELSEVIER

The most frequent phrasal verbs in English language EU documents – A corpus-based analysis and its implications

Anna Trebits

Eötvös Loránd University, Department of English Applied Linguistics, Language Pedagogy PhD Program, Budapest, Hungary

Abstract

This study explores the use of phrasal verbs in English language documents of the European Union (EU) as part of a larger-scale project examining the use of English in EU texts from various aspects including lexical, lexico-grammatical and textual features. Phrasal verbs, known to represent one of the most difficult aspects of learning English, are highly productive and widely used by native speakers. The purpose of this study is to identify the most frequent phrasal verb combinations in EU documents. To this end, an EU English Corpus of approximately 200,000 running words was built using texts which are representative of the fields of activities of the EU. The analysis revealed that the top 25 phrasal verbs account for more than 60% of all phrasal verb constructions in the corpus. The results also show that in terms of the frequency of phrasal verbs, EU documents show some similarity to written academic English. The paper also illustrates some instructional activities and the pedagogical relevance of the findings.
© 2009 Elsevier Ltd. All rights reserved.

Keywords: Phrasal verbs; Corpus analysis; EU documents; Course and materials design; English for Specific Purposes

1. Introduction

One of the most challenging aspects of learning English vocabulary is developing native-like proficiency of phrasal verbs in both spoken and written language (for example Laufer and Eliasson, 1993; Siyanova and Schmitt, 2007). Phrasal verbs, similarly to other multiword units such as multi-word verbs, idioms and collocations often pose problems even for advanced learners, especially those who learn English in a foreign language learning (EFL) situation in the classroom (Wray, 1999). The reasons for the notorious difficulty of mastering native-like use of phrasal verbs are manifold. Descriptive grammar books and research into avoidance in language learning suggest that some of the most decisive factors that may cause problems involve (1) the structural, grammatical features of phrasal verbs (e.g. word order), (2) semantic features (e.g. the degree of transparency of meaning), or (3) the fact that phrasal verbs may not exist in the learner's first language and therefore, they lack adequate strategies to recognize and process them (Biber et al., 1999; Liao and

E-mail address: a.trebits@yahoo.com

Fukuya, 2004). This problem is of much concern to language learners and teachers for the simple reason that phrasal verbs are a common feature of the English language. Although the problematic nature of phrasal verbs and other multiword units in English has been demonstrated by several recent studies (Gardner and Davies, 2007; Gilmore, 2004; Liu, 2003; Nesselhauf, 2003), it has also been generally suggested in most of them that language teaching and materials development does not appear to be informed by the research findings and that decisions about what to include in the syllabus "have usually been based on the author's gut-level impressions and anecdotal evidence of how speakers and writers use the language" (Biber and Reppen, 2002, p. 200).

As pointed out in many of the above cited papers, selecting the elements of a language course should instead be based on empirical corpus-based research findings. Using language data from a corpus, i.e. a carefully compiled collection of naturally occurring language (Biber et al., 1998) can prove to be useful for teaching both general spoken and written English, as well as specific registers usually referred to as English for Specific Purposes (ESP), for example business, legal, medical, engineering, etc. The relevance of the use of specially designed corpora built by language professionals to suit the specific needs of their students and/or research purposes in developing teaching materials for ESP is widely recognized in the literature (see Mudraya, 2006 for a review).

The present study is part of a larger-scale research project investigating the use of English in EU documents. The specific purpose of the project is to generate findings that may serve as a basis for the compilation of teaching materials and syllabi designed to develop the necessary language skills of those who work with English language EU documents, wish to be recruited to work for an EU institution, or major in EU, international relations or diplomacy studies. Previous studies on lexical and textual features of EU documents revealed the most frequent lexical elements and their collocation patterns (Trebits, 2008) and some aspects of the textual features exhibited by EU documents through the use of conjunctive cohesion. The aim of the current study is to explore the use of phrasal verbs in EU documents based on a corpus of EU texts representative of the various fields of activity the EU is currently involved in. A list of the most frequent phrasal verbs in the corpus will hopefully inform and productively contribute to the process of designing and writing materials for teaching 'EU English', that is English as it is used in EU documents (see Section 2.3).

2. Literature review

In this section of the paper I will outline the most important considerations regarding the definition of phrasal verbs as well as describe the fields of research that constitute the background to this study, namely corpus-based research on phrasal verbs and research focusing on the characteristics of English as it is used in EU documents.

2.1. Defining phrasal verbs

Phrasal verbs (PV) form part of a larger class of verbs usually referred to as multi-word verbs in comprehensive grammar books (Biber et al., 1999; Quirk et al., 1985) and in empirical research papers (for example Siyanova and Schmitt, 2007). Biber et al. (1999) distinguish between phrasal verbs, prepositional verbs, phrasal–prepositional verbs, free combinations and other multi-word verb constructions (of various patterns such as *take a look at, make fun of*, etc.) and discuss several structural and semantic criteria that can be used to differentiate between the various types. Despite the detailed descriptions and numerous examples provided for each category, the authors of the above-cited grammar books acknowledge the difficulty in making clear-cut distinctions between multi-word verbs, as many of them may belong to more than one category depending on the context. For example: *come back* may be interpreted either as a phrasal verb meaning "to resume an activity" or as a free combination meaning "to return" (Biber et al., 1999, point 5.3.1.1).

As Gardner and Davies (2007) point out, however, these purely grammatical distinctions, which may even be the subject of debates among linguists and grammarians, are of little instructional value for non-native speakers and learners of English struggling to master this area of vocabulary and grammar. They propose a more functional definition of phrasal verbs which states that two-word items consisting of a lexical verb and an adverbial particle are to be considered as phrasal verbs (PVs) in their study. For the purpose of the

present study, I decided to adopt the above definition, as, similarly to their research on phrasal verbs in the British National Corpus (BNC), the aim of the present study is to identify the most frequent PVs in a special corpus with a view to using the findings in materials design for language learners. The main purpose of this study is to inform the process of materials design so that learners of EU English have an opportunity to explore the use and meanings of those particular phrasal verbs which they are most likely to encounter in EU documents. Therefore, it was decided that no further grammatical distinctions will be made to differentiate between the verb + adverbial particle constructions revealed by the analysis.

2.2. Corpus-based research on phrasal verbs

In their recent study cited above, Gardner and Davies (2007) establish the list of the 100 most frequent PVs in the BNC. Their research aims, focusing on the pedagogical implications and applications of their findings, are fundamentally similar to the purpose of my study which is centred on a considerably smaller register, written English as it is used in EU documents, but which also hopes to contribute to teaching phrasal verbs and materials development in the field of ESP.

Previous corpus-based research on phrasal verbs include major phrasal verb dictionary projects which also helped produce a number of practice books for learners of English. For example, major ELT publishers such as Longman, Cambridge University Press or Oxford University Press have all published several reference and practice books based on corpus findings (for example English Phrasal Verbs in Use by McCarthy and O'Dell (1998) and Longman Phrasal Verbs Dictionary (2000) etc.). The comprehensive Longman Grammar of Spoken and Written English (Biber et al., 1999) and its abridged Student Grammar Version (Biber et al., 2002) both take a discourse perspective on the grammatical issues they address including phrasal verbs, and they are both based on corpus findings which are presented to the learner in detail.

Most research projects referred to above focus on general English, and less attention has been given so far to researching the use of phrasal verbs in the different registers of English, while many ESP learners could benefit from the findings of such research. My study is motivated by the fact that, to my knowledge, there have not yet been attempts to identify the most prominent phrasal verbs in EU documents using corpus linguistic methods, although such analyses could prove to be useful for learners of English working with EU documents.

2.3. Research on the use of English in the European context

The spread of English as the most important language of communication in international and European business organizations (Rogerson-Revell, 2006) and within EU institutions (Truchot, 2002), has warranted a number of diverse research projects mainly in translation studies and applied linguistics (for example: Jablonkai, 2008; McArthur, 2003; Modiano, 2003; Pym, 2000; Rogerson-Revell, 2006; Seidlhofer, 2007; Trebits, 2008; Trosborg, 1997; Truchot, 2002). The term 'Euro-English' denotes an emerging variety of English: English as a lingua franca in Europe used as a means of communication between native speakers of other languages (McArthur, 2003; Modiano, 2003; Seidlhofer, 2007). As Modiano (2003, p. 36) points out, Euro-English brings non-native speakers of English into contact with other non-native speakers in Europe "equally if not more often than with the native-speaker collective". If Euro-English is to be acknowledged as a variety of English similarly to Indian English for example, its lexical, lexico-grammatical, textual and even phonological characteristics will need to be described and standardized (Jenkins, 2001; Seidlhofer, 2001). Seidlhofer (2007) reports on a research project, which aims to compile a corpus of spoken English as a lingua franca (ELF) in Europe (called the Vienna-Oxford International Corpus of English (VOICE)) that could be used to carry out future research projects to describe the common lexico-grammatical and discourse features of spoken ELF use.

Another line of research investigates the use of English in the documents of the European institutions. The term 'EU English' is often used in these studies to refer to English as it is used in the documents of the institutions of the European Union (for example Jablonkai, 2008). These studies (e.g., Jablonkai, 2008; Trebits, 2008; Trosborg, 1997; Károly, 2007) often conclude that EU English displays a number of language features from various registers (for example: Legal English, Business English, Academic English) reflecting the diverse fields of activities the EU is actively involved in from commerce to educational programs. Therefore, research

on language features that distinguish EU English from other registers would benefit professionals from a wide range of different fields. As opposed to Euro-English, a non-native variety of English, EU English is produced by native speakers of English. Even if a particular text is not worded in English originally, it is translated into English by translators whose mother tongue (L1) is English, with the translations also being proofread and edited by L1 English speakers. Research projects focusing on EU English have investigated its lexical features with the help of corpora of EU documents (Jablonkai, 2008; Trebits, 2008), issues in terminology (Fischer, 2006) and translation (Trosborg, 1997).

3. Research questions

This study intends to point out the most frequent phrasal verbs in English language EU documents with the specific aim to assist the teaching of EU English by discussing the implications of the findings for instructional purposes. The research aims are as follows:

- To find the lexical verbs used in phrasal verb constructions.
- To identify the most frequent phrasal verbs in the Corpus.
- To determine the number of word-senses associated with the most frequent phrasal verbs occurring in the Corpus.
- To present the pedagogical relevance of the findings and suggest activity-types to teach phrasal verbs.

In what follows, I will describe the Corpus of EU English, present the analytical tools and discuss the analytical procedures used in this study. Then the presentation and the discussion of the findings will follow. The study concludes by relating the results to ESP research in general and to language pedagogy in particular.

4. Methodology[1]

4.1. The Corpus of EU English (CEUE)

In view of my focus on EU English for learners of English coming from a wide range of different professional backgrounds (e.g., trade, law, education, finance, etc.) it was important for the Corpus to reflect the diverse fields of activities the EU is involved in but not to be biased towards any of those fields in particular, so that the findings should remain balanced and pedagogically useful. Bearing that in mind, the following texts and documents were used to build the corpus:

- Nineteen information booklets (of about 20 pages on average) on the different activities of the EU (e.g., competition, market, economy, travelling, environment, justice, science, transport, etc.) intended for a general but informed audience. The booklets are downloadable from the internet at: http://ec.europa.int.,
- The annual general report on the activities of the EU in 2006 (published in early 2007 and available on the internet at: http://ec.europa.int.,
- Sample test material from recruitment competitions in all subject areas (the sample tests are available from the European Personnel Selection Office (EPSO) website and are freely downloadable).

The last EPSO sample test component of the CEUE was included because those who seek a job at an EU institution have to pass the EPSO recruitment competition which is not so much of a professional challenge for the candidates, but it definitely puts their knowledge of EU English to the test. The EPSO tests are convenient for the purpose of exemplifying how English is used in EU documents because they include texts on the most important subject matters (economy, law, education, etc.), as well as texts on general knowledge and understanding of EU institutions and tests of verbal reasoning. It is also important to note that all the texts included in the corpus were published after the year 2000.

---

[1] The method 4 of this paper is a revised version of the same section of the following article: Trebits, 2008.

The final version of the CEUE contains around 200,000 running words, and it consists of a balanced selection of texts in terms of the EU's fields of activity. As this study focuses on one single register, the language of EU documents and within that only written texts, the size of the corpus seems to be adequate for its purposes. Table 1 below summarizes the details of the Corpus.

4.2. Tools of analysis

I used the Lexical Frequency and Range computer programs by Heatley et al. (2002) to establish the frequency list of the CEUE and to compare it to the list of the most frequent adverbial particles and lexical verb + adverbial particle constructions in the written part of British National Corpus (BNC Written) comprising 89,800,000 tokens based on Leech et al. (2001). The lists obtained allowed me to draw up the list of the most frequent adverbial particles in the Corpus of EU English.

In order to examine the most frequently occurring lexical verbs which form verb + adverbial particle constructions in the CEUE, I used the concordance function of the Wordsmith Version 2 computer software (Scott, 1996). As concordance lists show the search word in context and provide frequency information too (Flowerdew, 1993), this function also allowed me to cross-examine the frequency lists produced by the Lexical Frequency Program. With the help of the concordancing function of the Wordsmith Tools software, all the verb + adverbial particle combinations were identified.

Finally, as one of the aims of this research concerned examining the number of word-senses associated with phrasal verbs in the CEUE, WordNet Version 2 (Miller, 2003) was also run on the most frequent verb + adverbial particle constructions.

4.3. Procedures of analysis

The main stages in identifying the most frequent phrasal verbs in the CEUE included listing all adverbial particles using comprehensive grammar books (Biber et al., 1999; Quirk et al., 1985) and conducting frequency and concordance analyses on the corpus. These analyses allowed me (1) to determine the number of times a particle-type functions as an adverbial particle as opposed to a preposition or other grammatical devices (e.g., *back* can act both as a noun and an adverbial particle, or *down* can both be a preposition and an adverbial particle in a sentence) and (2) to find the lexical verbs that form verb + particle constructions with them in the Corpus (e.g., *get back, break down,* etc.). The concordance lists produced all the inflectional forms of the same verb (e.g., *set, sets, setting* for the verb SET) which were then counted together. In order to ensure the reliability of the counts, both intra-rater and inter-rater reliability was used. Knowing that some phrasal verbs may have multiple meanings depending on the context, the next stage involved conducting semantic analyses on the most frequent phrasal verbs in the CEUE. The results of these analyses were compared to the number of word-senses associated with the same phrasal verbs in general English with the help of

Table 1
Summary of the contents of the Corpus of EU English.

Origin	Topics covered	Years of publication	Total number of words
Nineteen information booklets on the activities of the EU	Audiovisual arts, budget, consumer interests, customs policy, economy, enlargement, environment, foreign policy, globalization and trade, information society, justice, market, science, transport, travelling, the story of the EU and its institutions, general information on the EU	2002–2007	121,149
Annual general report on the activities of the EU	Agriculture, budget, customs, economy, education, enlargement, environment, globalization and trade, humanitarian aid, information society, justice, market, mobility, transport, overview of the activities of the institutions, police	2006	81,979
EPSO sample tests	General understanding of EU institutions, verbal and numerical reasoning, audit, law, economics	2006	3629
Total			206,757

Table 2
Lexical verbs used in verb + AP constructions in the CEUE.

ADV.P. Verbs	about	across	along	around	aside	away	back	behind	by	down	forward	in	off	on	out	over	under	up	Total
AGREE														11					11
BASE				1										70					71
BREAK						1				12			1					1	15
BRING	10						1			2				1	1			2	17
BUILD														12				3	15
CALL														15					15
CARRY														1	69	1			71
COME	2	1					1			3							4	2	13
DEPEND														28					28
DRAW														6				44	50
FIND															16				16
FOCUS														40					40
FOLLOW														3				13	16
GO	2						5		1	3	1			10					22
LAY										40									40
MAKE																		21	21
MOVE		3		14		1	1				6			1				1	27
OPEN																		31	31
POINT															10				10
PUT							1				32	1		3	1			1	39
REPORT															17				17
SET					3										72			85	160
STEP										1	1	1						8	11
TAKE						1	1				3		1	7	4	4		14	35
WORK														13	2				15

The verbs highlighted in green form verb + AP constructions with at least four different APs.

WordNet (Miller, 2003), on the basis of the Longman Phrasal Verbs Dictionary (2000) and the BNC Written (Leech et al., 2001). Given that the main goal of this research project is to contribute to developing practical ESP instructional materials, it was important to recognize that the same phrasal verb construction may have different meanings depending on the context, and to point out possible differences and similarities between general and EU English.

5. Results and discussion

5.1. The most frequent phrasal verb constructions in the CEUE

Table 2 presents the list of those lexical verbs and adverbial particles (APs) which proved to be the most productive in forming verb + particle constructions in the CEUE. A total of 130 lexical verb-types were identified in 187 verb + AP types (e.g., *point + out*, *cut + down*, etc.) and in a total of 1031 phrasal verb constructions. Considering the overall size of the corpus (about 200,000 tokens), these numbers become more meaningful knowing that this means that a learner of English will find one phrasal verb construction in approximately every 200 words of text. Since one page of typed text contains around 400–500 words, it can be assumed that reliably understanding phrasal verb structures is essential for those who work with EU documents on a regular basis. This finding lends further empirical support to the necessity of describing phrasal verbs in EU English in an effort to contribute to the development of language courses and teaching materials focusing on this particular register (Table 2).

It is interesting to note that 11 of the top 50 lexical verbs in the CEUE (e.g., *base, bring, call, go, make, open, put, report, set, take, work*) also function in PV constructions; moreover, they are among the top 25 phrasal verbs in the corpus (see Table 3 for details of the most frequent phrasal verbs). Some of the above verbs are also those which, besides being used frequently in PV structures, combine with the largest number of adverbial particles. For example, the verb *TAKE* combines with eight different particles in the CEUE to form phrasal

Table 3
The top 25 phrasal verbs in the CEUE.

Rank	Verb	Adverbial particle	Raw frequency	% Of all PVs in the CEUE
1	set	up	85	8.24
2	set	out	72	6.98
3	base	on	70	6.78
4	carry	out	69	6.69
5	draw	up	44	4.26
6	focus	on	40	3.88
7	lay	down	40	3.88
8	put	forward	32	3.10
9	open	up	31	3.01
10	depend	on	28	2.72
11	make	up	21	2.03
12	report	on	17	1.65
13	find	out	16	1.55
14	call	on	15	1.45
15	move	around	14	1.35
16	take	up	14	1.35
17	follow	up	13	1.26
18	work	on	13	1.26
19	break	down	12	1.16
20	build	on	12	1.16
21	agree	on	11	1.06
22	bring	about	10	0.97
23	go	on	10	0.97
24	point	out	10	0.97
25	speed	up	9	0.87
		Total	**708**	**68.67%**

verbs such as *take away, take back, take forward, take off, take on, take out, take over, take up*. Other prolific verbs in the CEUE include *break, bring, come, go, move, put, step*.

Table 3 below lists the 25 most frequent phrasal verbs in the CEUE. Interestingly, about half of the verbs on the list are among the 20 most frequent lexical verbs forming phrasal verb combinations in the BNC Written (*set, carry, put, make, find, move, take, work, break, bring, go, point*). Of much concern to EU English course designers and materials developers are the verbs which emerge as frequent in the CEUE but not at all so in general English. Learners need to be given ample opportunities to study and practice these verbs and their PV combinations since they may not have encountered them so often in general English course books.

As can be seen from the data, the top 10 phrasal verbs in the CEUE account for over 50% of all PV combinations while the top 25 phrasal verbs make up more than 60% of all phrasal verbs in the EU corpus. Therefore, it seems justified that these items be systematically dealt with in language courses focusing on the language of EU texts.

5.2. Semantic analysis of the most frequent phrasal verbs

Phrasal verbs, as many other one-word lexical verbs, may have a number of different meanings depending on the context in which they are used. Corpus-based vocabulary studies often point out the necessity to include analyses focusing on the possible meanings of particular word forms (for example Gardner and Davies, 2007). The meaning or meanings of a word in a given register can be very important to characterize the register itself. Furthermore, they are essential for the language learner and the language teacher in making decisions about the course syllabus they intend to follow (Trebits, 2008). Table 4 provides information regarding the word-senses associated with the 25 most frequent phrasal verbs in the CEUE in comparison to word-senses of the same phrasal verbs in general English (BNC). This semantic analysis was carried out with the help of WordNet (Miller, 2003), the Longman Phrasal Verbs Dictionary (2000) and the written part of the BNC (Leech et al., 2001).

It is not surprising that EU English uses considerably fewer meanings of the phrasal verbs with multiple meanings than general English where the counts are based on both the spoken and written registers and several genres within those. As confirmed by corpus-based studies on lexico-grammatical features, written academic English uses substantially fewer phrasal verbs than other written genres (e.g., fiction, news) or spoken English (Biber et al., 1999). In this respect, written EU English seems to show much resemblance to academic writing. Nonetheless, about one third of the most frequent phrasal verbs are used to express more than one meaning in the CEUE. For example, the phrasal verb *set* up has two different meanings in the CEUE (1 start a business or organization, 2 make the arrangements for something to happen) and more than ten in general English. This finding underlines the importance of designing activities for language learners that allow them to explore this very important aspect of phrasal verbs (Table 4).

5.3. Corpus-based data-driven activities to teach EU English

The following activities are all based on language data from the CEUE. The merits of data-driven activities have been highlighted in a number of studies focusing on teaching English for both general and specific purposes (for example, Mishan, 2004; Mudraya, 2006; Trebits, 2008). Data-driven activities have the advantage of allowing students to access the real-life language use of their particular context of interest. The first stage of doing corpus-based data-driven activities involves presenting students with language data from the corpus, often in the form of selected lines from a concordance list, that is, a list of all the occurrences of a particular word in the corpus in context. As concordance lists tend to be extremely long, the teacher may choose to present a selection of lines from the full list to make sure that the amount of data is not overwhelming for the students and that the lines selected contain all the necessary information. The students' task is, then, to work out, for example, differences in meanings of the same word form, collocations and grammatical structures on the basis of the data. Working with language data not only gives students an opportunity to explore the structural and semantic aspects of language, but it also improves their language learning skills. Traditional practice exercises (e.g., gap-filling, paraphrasing) may follow this stage to ensure further practice and to prepare stu-

dents for the cognitively more complex tasks which, then, put the patterns they learned in a wider context. To sum up, the activities can be presented in the following order:

- Language data (selected concordance lines),
- Practice exercises (e.g., gap-filling, paraphrasing),
- Complex practice activities (usually involving several skills).

The sample activities below give an example of how language data from the CEUE may be exploited to teach phrasal verbs using a data-driven approach.

5.3.1. Sample activities to teach and practice the phrasal verbs set up and set out

The verb *set* combining with just two adverbial particles (*up, out*) accounts for more than 10% of all phrasal verb constructions in the CEUE; therefore, mastering the differences between these two PVs is extremely important for those who work with EU documents on a regular basis. The following sample tasks explore both the semantic and the grammatical aspect of these PVs.

Selected lines of the concordance list of the verb *set*.

1. The 'roadmap' of 15 July 2005, which **set out** possible future changes to European
2. The European Parliament **set up** a committee of inquiry to look into
3. These principles are **set out** in the EU's General Food Law of 2002,
4. Subsequent action plans have **set out** roadmaps of what needs to be done by
5. Many of the challenges **set out** in the May reports had been addressed
6. 22 February, the Commission **set out** its initial ideas on the structure and
7. Burden that the measures **set out** in the seven proposals might entail.
8. And the Netherlands have **set up** supermarkets all across central and
9. The European Commission was **set up** in the 1950s under the EU's founding
10. Of southeast Europe have **set up** a single energy community in which

Sample task 1: By looking at the concordance sample of the verb *set*, make a list of things that you can *set up* and things that you can *set out*. Compare your data with information in a monolingual dictionary. Summarize your findings in the following table. (Some possible answers are given).

	CEUE sample	Longman PV dictionary
set out	*future changes*	*guidelines*
set up	*a committee*	*a business*

Sample task 2: Based on the information you found, match the definitions with the correct phrasal verb.

(a) `To make arrangements for something to happen.`
(b) `To explain ideas in a clearly organized way.`

Table 4
Semantic analysis of the most frequent phrasal verbs in the CEUE.

Top 25 phrasal verbs in the CEUE	# Of word-senses in the CEUE	# Of word-senses in WordNet and LDPV
set up	2	15
set out	1	3
base on	1	1
carry out	1	2
draw up	1	5
focus on	1	1
lay down	1	2
put forward	2	4
open up	2	7
depend on	2	3
make up	1	9
report on	1	1
find out	1	4
call on	1	1
move around	1	2
take up	2	13
follow up	1	2
work on	1	2
break down	4	8
build on	1	1
agree on	1	1
bring about	1	2
go on	2	5
point out	1	3
speed up	1	2

Words highlighted in orange have multiple meanings in both corpora.

Sample task 3: Complete the following sentences using the words *up* or *out*.

```
(1) The details of the proposal are set ___ in this booklet.
(2) After finishing his studies, he set ___ a business, but it failed.
(3) Handouts should set ___ information in a reader-friendly format.
(4) Self-study centres will be set ___ at the university.
```

Sample task 4: By studying the concordance lines above, make a list of who or what can *set up* or *set out* something by filling in the following table.

the roadmap		*future changes*
	set out	
	set up	

Sample task 5: Find all the passive constructions in the concordance sample. What do they tell you about the third form of the verb *set*?

Sample task 6: The phrasal verb *set up* has more than 10 different meanings in general English. Use a phrasal verbs dictionary and compare the meaning(s) of the verb in the CEUE and in general English.

6. Conclusion

The main aim of this study was to contribute to syllabus and materials design for teaching EU English by revealing the most frequently used phrasal verbs in the Corpus of EU English, a collection of texts published by European institutions. The study intended to answer the question of which phrasal verbs to include in an EU English syllabus, assuming that selecting items of such an important part of English lexis and grammar cannot be made solely on language teachers' and linguists' intuition, however experienced they may be in this field (Biber et al., 1999; Biber and Reppen, 2002; Gardner and Davies, 2007). The findings of the study may serve as a starting point in developing the phrasal verb component of a course focusing on the language of EU documents. Further analyses of concordances of less frequent phrasal verbs would help determine other important PV items to include in the syllabus. The study also provides examples of activity-types based on language data from the CEUE.

An obvious limitation of the study is the size of the corpus, since a larger corpus would allow more precise observations to be made. Previous studies (e.g., Trebits, 2008) involving research on the lexical and textual features of EU texts will be complemented by further investigations as regards other lexico-grammatical features of language use in a larger corpus of EU English. It is hoped that these studies will make a meaningful contribution to the design of courses materials focusing on the use of the English language in the documents of the European Union.

Acknowledgements

References

Biber, D., Conrad, S., Reppen, R., 1998. Corpus Linguistics. Cambridge University Press, Cambridge.

Biber, D., Johansson, S., Leech, G., Conrad, S., Finegan, E., 1999. Longman Grammar of Spoken and Written English. Longman, Harlow.

Biber, D., Conrad, S., Leech, G., 2002. Longman Student Grammar of Spoken and Written English. Longman, Harlow.

Biber, D., Reppen, R., 2002. What does frequency have to do with grammar teaching? SSLA 24, 199–208.

Fischer, M., 2006. Translation (policy) and Terminology in the European Union. Paper Presented at the International Conference on Terminology, Lessius University College, Antwerpen.

Flowerdew, J., 1993. Concordancing as a tool in course design. System 21, 231–244.

Gardner, D., Davies, M., 2007. Pointing out frequent phrasal verbs: a corpus-based analysis. TESOL Quarterly 41, 339–359.

Gilmore, A., 2004. A comparison of textbook and authentic interactions. ELT Journal 58, 363–374.

Heatley, A., Nation, I.S.P., Coxhead, A., 2002. Range and Frequency Programs Software. Available from: <http://www.vuw.ac.nz/lals/staff/Paul_Nation>.

Jablonkai, R., 2008. Defining EU Discourse. Steps Towards the EU English Discourse Corpus. Paper Presented at the SEAS PhD Conference, Eötvös Loránd University, Budapest.

Jenkins, J., 2001. Euro-English accents. English Today 68 (17), 16–19.

Károly, K., 2007. Szövegtan és fordítás [Text Linguistics and Translation]. Akadémiai Kiadó, Budapest.

Leech, G., Rayson, P., Wilson, A., 2001. Word Frequencies in Written and Spoken English. Longman, London.

Laufer, B., Eliasson, S., 1993. What causes avoidance in L2 learning: L1–L2 difference, L1–L2 similarity or L2 complexity? SSLA 15, 35–48.

Liao, Y.D., Fukuya, Y.J., 2004. Avoidance of phrasal verbs: the case of Chinese learners of English. Language Learning 54, 193–226.

Liu, D., 2003. The most frequently used spoken American English idioms: a corpus analysis and its implications. TESOL Quarterly 37, 671–700.

Longman Phrasal Verbs Dictionary, 2000. Pearson Education Limited, Harlow.

McArthur, T., 2003. World English, Euro-English, Nordic English? English Today 73, 54–58.

McCarthy, M., O'Dell, F., 1998. English Phrasal Verbs in Use: Intermediate. Cambridge University Press, Cambridge.

Miller, G., 2003. WordNet (Version 2.0) Software. Available from: <http://www.cogsci.princeton.edu/wn/obtain.shtml>.

Mishan, F., 2004. Authenticating corpora for language learning: a problem and its resolution. ELT Journal 58 (3), 219–227.

Modiano, M., 2003. Euro-English: a Swedish perspective. English Today 74 (19), 35–41.

Mudraya, O., 2006. Engineering English: a lexical frequency instructional model. English for Specific Purposes 25, 235–256.

Nesselhauf, N., 2003. The use of collocations by advanced learners of English and some implications for teaching. Applied Linguistics 24, 223–242.

Pym, A., 2000. The European Union and its future languages: questions for language policies and translation theories. Across Languages and Cultures 1, 1–17.

Quirk, R., Greenbaum, S., Leech, G., Svartvik, J., 1985. A Comprehensive Grammar of the English Language. Longman, New York.

Rogerson-Revell, P., 2006. Using English for international business: a European case study. English for Specific Purposes 26, 103–120.

Seidlhofer, B., 2001. Towards making 'Euro-English' a linguistic reality. English Today 68 (17), 14–16.

Seidlhofer, B., 2007. Common property: English as a lingua franca in Europe. In: Cummins, J., Davison, C. (Eds.), International Handbook of English Language Teaching. Springer, New York, pp. 137–153.

Scott, 1996. Wordsmith Lexical Analysis Tools Software. Oxford University Press, Oxford. Available from: <http://www.lexically.net/wordsmith/>.

Siyanova, A., Schmitt, N., 2007. Native and nonnative use of multiword vs. one-word verbs. IRAL 45, 119–139.

Trebits, A., 2008. English Lexis in the Documents of the EU – A Corpus-based Exploratory Study. Eötvös Loránd University, Budapest: Working Papers in Language Pedagogy 2. Available from: <http://langped.elte.hu/Wopalpindex.htm>.

Trosborg, A., 1997. Translating hybrid political texts. In: Trosborg, A. (Ed.), Text Typology and Translation. John Benjamins, Amsterdam, pp. 145–159.

Truchot, C., 2002. Key Aspects of the Use of English in Europe. Council of Europe, Strasbourg.

Wray, A., 1999. Formulaic language in learners and native speakers. Language Teaching 32, 213–231.