



This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

<http://www.elsevier.com/copyright>



Conjunctive cohesion in English language EU documents – A corpus-based analysis and its implications

Anna Trebits *

Department of Applied Linguistics, Language Pedagogy PhD Program, Eötvös Loránd University, Rákóczi út 5, H-1085 Budapest, Hungary

Abstract

This paper reports the findings of a study which forms part of a larger-scale research project investigating the use of English in the documents of the European Union (EU). The documents of the EU show various features of texts written for legal, business and other specific purposes. Moreover, the translation services of the EU institutions often produce texts that exhibit features which in terms of textual organization neither link them to the source language nor to the target language of texts. The aim of the present corpus-based study is to describe one type of cohesion: the use of conjunctions in EU documents in order to uncover some of the textual organization patterns they show. Therefore, an EU English Corpus of approximately 200,000 running words was built using texts which represent the diverse fields of activities of the EU. The analysis compares the use of conjunctions in EU-related and general English texts using the database of the British National Corpus (BNC). The paper also illustrates some data-driven instructional activities that may be used in EFL/ESL classrooms when teaching English for EU purposes.

© 2009 The American University. Published by Elsevier Ltd. All rights reserved.

1. Introduction

The use of English as a lingua franca both in the international and in the European contexts is well-established and has been the object of a number of diverse research projects in translation studies and applied linguistics (e.g., McArthur, 2003; Modiano, 2003; Pym, 2000; Rogerson-Revell, 2006; Trosborg, 1997; Truchot, 2002). Although the EU itself regularly issues policy statements that aim to help advertise and preserve linguistic diversity in Europe, there is a growing body of research that demonstrates the spread of English in both European business organizations (Nickerson, Gerritsen, & van Meurs, 2005; Rogerson-Revell, 2006) and within the EU institutions as well (Truchot, 2002). The most recent enlargements of the European Union entailed that twelve countries which predominantly use English in their international communication (Truchot, 2002) joined the EU, which will probably reinforce the further establishment of the English language in European communication.

* Tel./fax: +36 1 239 1545.

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

The current spread of English in Europe is of much importance to English language teachers because it entails that there is also a heightened interest among language students in studying *EU English*, that is, learning the characteristics of English as it is used in EU documents and EU institutions. This is especially true in countries where the national and/or official language does not belong to one of the major working languages of the EU (like French or German). *EU English* is gaining importance within the field of English for Specific Purposes (ESP) as new member states prepare for their turn at the EU Council Presidency. To date, most of the recent publications intended to help speakers and students of English familiarize themselves with the linguistic features of English language EU documents approached this issue from a translation point of view or focused on specialized technical vocabulary (e.g., [EU ABC online](#); [Jablonkai, Erdei, Fekete, Homolya, & Nagy, 2006](#); [Klaudy & Bart, 2003](#)). Yet, many of those who work with EU documents in their home country or wish to be recruited to an EU institution do not translate but use the information in the documents. Consequently, they need to be proficient enough to understand the documents reliably, and to produce other EU-related texts complying with the EU norms. To my knowledge, comprehensive monolingual course books written for students of English focusing on the typical features of EU English and based on corpus analyses of EU documents are not yet available.

The present study forms part of a larger-scale research project which investigates the use of English in EU documents using corpus linguistic techniques with the specific aim to serve as a basis for the compilation of specialized teaching materials designed to develop the necessary EU English language skills. A previous study on English lexis in EU documents ([Trebits, 2008](#)) uncovered the most frequent technical and non-technical lexical elements and collocation patterns that would need to be brought to students' attention and included in an EU English course. The present study focuses on the use of conjunctions in EU documents. Conjunctions are important cohesive devices and the mastery of their use has a huge impact on students' textual competence ([Bachman, 1990](#)) or discourse competence ([Canale & Swain, 1980](#) cited in [Bachman, 1990](#)), one of the major constituents of communicative competence. This study is motivated by the fact that little research has been done on the use of conjunctions in EU documents from an ELT viewpoint even though many language learners who need to work with EU documents on a daily basis both inside and outside the EU institutions could benefit from the findings of such research.

2. Background to the study

2.1. *English in Europe and in EU documents*

In order to situate this study in the context of ESP, it is important to define the terms 'Euro-English' and 'EU English', both referring to varieties of English in the European context (see e.g., [Jablonkai, 2008](#); [Modiano, 2001](#); [Modiano, 2003](#)). The English language is recognized today as "the most universal linguistic entity that humankind has ever known" ([McArthur, 1998](#), p. 57) with many existing varieties across the globe. The term 'Euro-English' refers to a relatively new variety of English: English as a lingua franca in Europe used by the linguistically and culturally diverse peoples of Europe ([McArthur, 2003](#); [Modiano, 2003](#); [Seidlhofer, 2001](#)). 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". An obvious consequence of this phenomenon is the emergence of typical lexical choices, discourse strategies and even accents ([Jenkins, 2001](#)) that will need to be described and codified if Euro-English is indeed to be acknowledged as a "linguistic reality" similarly to other varieties of English such as Indian English for example ([Seidlhofer, 2001](#), p. 14). [Seidlhofer \(2001\)](#) and [Jenkins \(2001\)](#) report on recent research efforts to describe and standardize this emerging variety. The Vienna-Oxford ELF Corpus project (VOICE), for example, aims to compile a corpus of spoken English as a lingua franca in Europe (ELF) to be used to describe the common lexico-grammatical and discourse features of spoken ELF use ([Seidlhofer, 2005](#); [Seidlhofer, 2007](#)).

The term 'EU English' refers to English as it is used in the documents of the institutions of the European Union (e.g., [Jablonkai, 2008](#); [Trebits, 2008](#)). The documents of the EU show various features of texts written for specific purposes (e.g., legal, medical, business) mirroring the wide range of activities the EU is involved in. Moreover, considering the characteristics [Swales \(1990\)](#) proposes for the identification of a discourse community – including an agreed set of common public goals, a mechanism of intercommunication between

members, the existence of discursual expectations as to the genre utilized for intercommunication and members possessing a minimum degree of expertise regarding content and discourse – the EU institutions may be collectively regarded as one discourse community imposing its own set of norms on its members. Studies examining the language issues and translation practices of the EU institutions conclude that in an effort to minimize translation problems, EU texts are standardized in terms of terminology, lexis and structural organization (Károly, 2007; Pym, 2000; Trosborg, 1997). Contrary to Euro-English, which is a non-native variety of English, EU English is produced by native speakers of English. When a particular text is originally worded in a different working language of the EU, it is translated into English by native speakers of English as translators always translate into their mother tongue (L1) in the EU institutions. Research projects focusing on EU English have investigated lexical features with the help of corpora of EU documents (Jablonkai, 2008; Trebits, 2008), issues in terminology (Fischer, 2006) and translation (Trosborg, 1997).

2.2. Textual competence, cohesion and conjunctions

Knowing a language involves a wide range of skills and competences whose mastery contributes to one's overall proficiency. Several models have been proposed to describe the elements of communicative competence (see Bachman, 1990 for a review), all of which include the skill of constructing a well-organized cohesive and coherent text referred to as discourse competence in Canale and Swain (1980) or textual competence in Bachman (1990). A quick survey of ELT materials reveals that discourse/textual competence is a complex skill which is given some attention in most second or foreign language courses (ESL/EFL) and more emphasis in ESP courses where it is dealt with explicitly and in more detail (e.g., writing a report, writing a formal letter, making presentations).

Making language students aware of the fact that discourse/textual skills need to be learned and practiced systematically is especially important in the light of the growing body of research on cultural differences in rhetorical conventions (e.g., Vergaro, 2004) suggesting that “language and writing are cultural phenomena” (Bereczky, 2007, p. 81), and thus different conventions and norms apply in different languages. Contrastive rhetoric studies comparing various genres produced in different languages (e.g., Johns, 1980) conclude that languages have their own rhetoric conventions, “which may reflect different writing conventions that are learned in a culture” (Connor, 1996, p. 16). Consequently, second/foreign language (L2) learners should be aware that without conscious attention to and practice of discourse/textual skills, L1 interference in the texts they produce may influence the interpretation of these texts by L2 readers. In the context of multinational EU institutions, readers may have very different linguistic backgrounds (often neither the writer's L1 nor L2, the working language), which gives the mastery of discourse/textual competence even more importance.

For lines of words to be considered a text, a piece of writing must meet seven standards of textuality (de Beaugrande & Dressler, 1983 cited in Bereczky, 2007), otherwise the text cannot be considered communicative, and it cannot even be considered a text. According to de Beaugrande and Dressler (1983, cited in Bereczky, 2007) the seven standards of textuality are as follows: cohesion, coherence, intentionality, acceptability, informativity, situationality and intertextuality. In their definition cohesion:

concerns the ways in which the components of a surface text ... are mutually connected within a sequence ... cohesion rests upon grammatical dependencies. All of the functions which can be used to signal relations among surface elements are included under our notion of cohesion. (p. 3)

In other words, cohesion is a pattern of relations between lexical elements and structures like words, phrases and sentences that are combined to build an integrated and logical text. The same authors list several grammatical devices by which cohesion can be reached, one of which is conjunction (some of the others include: tense, aspect, ellipsis, paraphrase). In their influential work *Cohesion in English*, Halliday and Hasan (1976) also include conjunctions on their list of cohesive devices. A conjunction is a type of function word that joins words, phrases or clauses together. As Halliday and Hasan (1976) point out, “conjunctive elements are cohesive not in themselves but indirectly, by virtue of their specific meanings” (p. 226) therefore, they consider that any expression of a semantic relation which can operate conjunctively falls within the category of conjunction, be it an adverb, compound adverb, prepositional phrase or linking adverbial (e.g., *furthermore*, *as a result*, *in addition to*, etc.). Many comprehensive grammar books (Biber, Johansson, Leech, Conrad, &

Finegan, 1999; Quirk, Greenbaum, Leech, & Svartvik, 1985) classify conjunctions and linking adverbials separately. According to their categorization, conjunctions may connect two clauses of equivalent syntactic status (coordinators) or they may link a main clause to a dependent clause (subordinators), while linking adverbials also function as important cohesive devices in that they link passages of text together at sentence boundaries. Conjunctions and linking adverbials often overlap (some conjunctions can not only connect clauses but sentences too (e.g. *so*, *yet*, *but*)).

As the purpose of this study is to describe conjunctive cohesion in English language EU documents with a focus on implications for language teaching, I decided to adopt the position of Halliday and Hasan (1976) as regarding what counts as a conjunction, and I will treat both conjunctions and sentence or linking adverbials as one category. Two main reasons motivated my choice: (1) both conjunctions and sentence adverbials are important cohesive devices, that is, their use or misuse can seriously affect the interpretation of a text, and (2) by combining the two categories there is a considerably wider choice of words to analyze which allows me to draw more precise conclusions regarding cohesive conjunction in English language EU documents. Table 1 summarizes the semantic categories of conjunctions based on the works of Quirk et al. (1985), Biber et al. (1999), and Halliday and Hasan (1976) and provides examples to illustrate each category. It is important to note the difference in the way the various categories are labeled by the authors cited above. When the same category appears under different names, all of these are included in the table and divided by a slash (/).

Thus the categories of conjunctions (including linking adverbials) examined in the present study are as follows: additive, adversative, causal, temporal, continuative, hypothetical, clarifying.

2.3. Corpus linguistics and data-driven language learning

One of the challenges language teachers – especially in a foreign language context (EFL) – must rise to is bringing real world language use closer to their learners. Corpus linguistic techniques and computer technology can provide significant help in pointing out ‘real life’ language use in context with the help of language corpora, that is, stretches of texts representative of particular registers or varieties of language (Biber, Conrad, & Reppen, 1998; Kennedy, 1998). Several corpora are made available on the internet (Mishan, 2004), but language professionals can also carefully build their own corpora based on their particular research purposes and/or the immediate needs of the students they teach. An array of corpus-based studies demonstrates the significance of these techniques for materials development for both general English and ESP courses. Many of them are focused on various aspects of vocabulary learning, for example, idioms, phrasal verbs, specialized

Table 1
Categories of conjunctions with illustrative examples.

Categories	Coordinators	Subordinating conjunctions ^b
Additive	<i>and, or, nor, both...and... either...or... neither...nor...</i>	<i>in addition, besides, moreover, furthermore, additionally, etc.</i>
Adversative/disjunctive/contrastive/concessive ^a	<i>but</i>	<i>though, although, while yet, however, nevertheless, nonetheless, at the same time</i>
Causal	–	<i>because, as, for, since, in view of, as a result, consequently, that's why, etc.</i>
Temporal	–	<i>after, as long as, until, after that, at the same time, meanwhile, next, when, etc.</i>
Continuative/transitive/changing subject	–	<i>anyway, now, regarding, as regards, with reference to, etc.</i>
Hypothetical/concessive ^c	–	<i>if, provided that, unless, in case, in that case, just in case, if so, if not, provided that, as though, etc.</i>
Clarifying/focusing/appositive	–	<i>for example, in particular, for instance, that is, actually, I mean, in other words, etc.</i>

^a Meaning opposition.

^b Also referred to as *sentence adverbials* or *adverbial phrases* in the literature.

^c Meaning condition.

registers: academic, EU, engineering, just to name a few (Cullen & Kuo, 2007; Gardner & Davies, 2007; Gilmore, 2004; Liu, 2003; Mudraya, 2006; Thurstun & Candlin, 1998; Trebits, 2008). Other studies and recently published grammars apply corpus-based analyses to the description and teaching of grammatical features (Biber & Reppen, 2002; Biber et al., 1999; Hughes & McCarthy, 1998).

Data-driven learning (DDL) is a corpus-based methodology which uses the merits of corpus linguistics to engage students in meaningful and autonomous language learning. Genuine discovery is at the heart of DDL where students use concordance lists to uncover some of the lexical, syntactic and even semantic properties of selected words (John's, DDL page; Mishan, 2004; Mudraya, 2006). Concordance and key words in context (KWIC) lists display several lines from a corpus, each line containing the chosen key word(s). Studying the lines and the neighboring words allows the learner to discover what company the word keeps, that is, which words occur most often together with the chosen word. This may lead to the discovery of some very important pieces of information about collocations, idioms, verb tenses, word order, style, just to mention a few. Considering the claims of the Involvement Load Hypothesis (Laufer & Hulstijn, 2001), according to which language learning is enhanced by the learners being engaged in various tasks and even in the process of discovering the meaning(s) of a word through inference for example, DDL activities may have great potential for language teaching.

3. Research aims

The aim of this study is to examine the use of conjunctions in EU documents in English. Conjunctions are used to express logical relations within the text and their use can influence how a text is interpreted to a considerable extent. Moreover, due in part to the so-called 'full stop rule'¹ observed by translators, readers of EU documents may often feel that the way conjunctions are used in them is somewhat different from the way they are used in national documents. The study also intends to reveal the implications of the findings for teaching English for EU Purposes. Therefore, the research questions are as follows:

- (1) What are the most frequently used conjunctions in the Corpus of EU English?
- (2) How does the frequency of conjunctions compare in the Corpus of EU English and the British National Corpus?
- (3) What are the most prevailing usage patterns of some of the most frequent conjunctions in the Corpus of EU English?
- (4) How can the Corpus of EU English be used to design data-driven learning activities?

In the following section of the paper, I will present the Corpus of EU English and discuss the analytical tools and procedures used in this study. Then I intend to present and interpret the results of the analyses, as well as to include some sample DDL activities that exemplify how to incorporate the development of discourse competence in language lessons. The study ends with a discussion of the relevance of the findings for teaching English for EU purposes.

4. Method²

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) it was important for the Corpus to reflect the diverse

¹ The 'full stop rule' prescribes that the sentence boundaries of the source text have to be kept in the translated versions of the same text for easier future references to particular sections or lines of the text.

² Section 4 of this paper is a revised version of the same section of the following article: Trebits (2008). English lexis in the documents of the EU – A corpus-based exploratory study. Eötvös Loránd University: Working Papers in Language Pedagogy, 2. Available online at: <http://lagped.elte.hu/Wopalpindex.htm>.

Table 2

Summary of the contents of the Corpus of EU English.

Origin	Topics covered	Years of publication	Total number of words
19 Information booklets on the activities of the EU	Audiovisual arts, budget, consumer interests, customs policy, economy, enlargement, environment, foreign policy, globalization & trade, information society, justice, market, science, transport, traveling, 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 & 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 & numerical reasoning, audit, law, economics	2006	3629
Total			206,757

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. Table 2 summarizes the details of the Corpus.

The last EPSO sample test component of the Corpus was included because those who seek a job at an EU institution have to pass the EPSO recruitment competition which definitely puts their knowledge of EU English to the test. The EPSO tests also comprise a considerable number of writing tasks (e.g., report writing, essay writing), so the mastery of discourse/textual skills is essential to pass. The EPSO tests are convenient for the purpose of exemplifying how English is used in EU documents because they also 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. All the texts included in the CEUE were published after the year 2000. 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 only on written EU documents, the size of the corpus seems to be adequate for its purposes.

4.2. Tools and procedures of analysis

I used the Lexical Frequency and Range computer programs by Heatley, Nation, and Coxhead (2002) to establish the frequency list of the Corpus and to compare it to the list of the most frequent conjunctions in the written part of British National Corpus (BNC Written) comprising of 89,800,000 tokens. The frequency lists obtained allowed me to: (1) draw up the list of the most frequent conjunctions in the Corpus of EU English and (2) compare the two corpora in terms of the categories of the conjunctions used.

In order to examine the most frequently occurring conjunctions in context and to reveal their most important usage patterns, I used the concordance function of the WordSmith computer software (Scott, 1996). As one of the aims of this study is to compare the use of conjunctions in general English and EU English, concordance lists from the BNC Written were also examined to compare some of the most frequent conjunctions in both corpora. Finally, concordance samples from the CEUE were selected to design activities that help students acquire some important discourse skills.

5. Results and discussion

This section of the paper presents the findings of the analyses by relating them directly to how they could be used as a basis for materials development in one area of ESP – English for EU purposes. In the ensuing discussion, I will point out the most important general characteristics of conjunctive cohesion in EU texts based on the CEUE, and I will present examples of the most interesting usage patterns revealed by the analyses. This section ends with the presentation of a series of DDL activities designed for the teaching of EU English.

5.1. Most frequent conjunctions in the corpus of EU English (CEUE)

The CEUE of about 200,000 tokens and over 9000 word-types was analyzed with the help of the WordSmith Tools software (Scott, 1996) and The Lexical Frequency and Range Programs (Heatley et al., 2002). The main stages in identifying the conjunctions in the corpus included listing all of the examples of conjunctions from comprehensive grammar books and handbooks on advanced writing used in advanced linguistics and academic writing courses at Eötvös University and using the list of conjunctions and adverbials from the BNC (Biber et al., 1999; BNC online; Leech, Rayson, & Wilson, 2001; Quirk et al., 1985), sorting them into the categories established based on the same sources (as discussed in Section 2.2), and conducting frequency analyses on the corpus by matching the list of conjunctions with the word frequency output list. As some conjunctive expressions have the same form as prepositions and conjunctions (e.g., *after*, *before*, *until*), in those cases the frequency data was double-checked with the help of concordances which show the words in context and allow for sorting out the number of occurrences in each function of the word. It was also important to take into account that some subordinators may have more than one semantic role, such as *as*, *since* or *while*. In an effort to ensure that this important aspect of the analysis was properly carried out, concordance lists were produced from the CEUE for each conjunction with multiple meanings. A close examination of the concordance lists allowed for sorting out how different meanings of the same conjunction-form were distributed in the CEUE. For example, a common subordinator *since* may be used both as a temporal and as a causal conjunction. It appears in the CEUE a total of 112 times, the vast majority of which (91 times) as a temporal subordinator, and only 21 times as a causal subordinator. An extract of the concordance list for *since* can be seen in Fig. 1. In lines 2, 5 and 12 *since* acts as a causal subordinator, while in the rest of the sample sentence extracts it is used as a temporal conjunction.

Another important step in finalizing frequency counts in the CEUE involved separating one-word conjunctions (e.g., *as*, *for*, *so*) from multiword conjunctions comprising the same word form (e.g., *as well as*, *as for*, *for example*, *so as to*, *so that*). For the counts to be reliable, it was essential to make sure that, for example, in the overall count of *as*, the occurrences of *as well as* were not included. To that end, the frequency counts of multiword conjunctions were deducted from the counts of the one-word conjunctions they contained (e.g., counts for *for example* were deducted from counts for *for*). In sum a total of 94 conjunction-types were identified occurring 12,197 times in the CEUE.

The word frequency analysis revealed that the number of conjunctions in the CEUE was generally comparable and proportionate to the register of academic prose in the BNC. According to Leech et al. (2001) academic prose in the BNC contains around 18,000 subordinators and about 42,000 coordinators per million words. As shown in Table 3 the CEUE presents similar figures, which is not surprising considering that it is comprised of written official and semi-official documents which are generally close to written academic English in style.

A more detailed analysis using the seven categories of conjunctions described in Section 2.2 shows that additives are by far the most frequent devices of conjunctive cohesion, followed by temporals and causals. Adversative, clarifying and hypothetical conjunctions were found to be much less frequent in the texts of the CEUE, and the least frequently used were continuatives. The prevalence of additives is not surprising

1	Rail has been in decline. Since 1970, its market share has fallen from 10 to 6
2	membership has risen from six to 27. Since the Union is open to any European country
3	signatories to the Cotonou Agreement. Since 2005, they have been able to export
4	further than earlier agreements, since it has moved from trade relations based on
5	environmental and regional policies. Since it was no longer a purely economic community,
6	to take stock of the progress made since the Guadalajara Summit in 2004
7	amendments to Commission proposals. Since September summaries of the Commission's
8	developments which had arisen since the initial 1998 recommendation
9	more than 400 000 young people since 1995. Exchanges are a pathway to a different
10	violence that has scarred the world since the 11 September 2001 attacks
11	beginning, not the end, of a process. Since then, the basic structure of the
12	EU's external borders. Moreover, since criminals can also exploit freedom of movement

Fig. 1. Concordance sample from the CEUE for the conjunction *since*.

Table 3

Coordinating and subordinating conjunctions in the CEUE.

Category	Number of occurrences	Normed count (per million words)
Coordinators	8418	42,090
Subordinating conjunctions	3846	19,320
Total	12,264	61,320

Table 4

The rank order of conjunction types in the CEUE.

Rank	Category	Total number of occurrences
1	Additive	8673
2	Temporal	1082
3	Causal	912
4	Adversative	548
5	Clarifying	500
6	Hypothetical	345
7	Continuative	204

Table 5

The distribution of coordinators in the CEUE and the BNC.

Coordinators	CEUE		BNC Written
	Raw counts	Normed counts (to 1 million words)	Normed counts (to 1 million words)
And	7375	36,875	27,002
Or	718	3590	3747
But	320	1600	4370
Nor	5	25	136

considering that many EU documents are usually descriptive and/or prescriptive in nature (e.g., official progress reports, country reports, recommendations, decisions). Many of the texts are also argumentative and wish to convince the reader by presenting arguments in support of the EU's position on different matters, hence the importance of temporals and causals, second and third in frequency ranking (see Table 4 for more details).

5.2. Conjunctions in the Corpus of EU English and the British National Corpus

Given the difference in the size of the two corpora, the frequency counts were normed to one million words in both so that the comparison would be feasible. To obtain the normed frequency counts of the BNC Written, two main sources (Biber et al., 1999; Leech et al., 2001) as well as the online version of the BNC were used. Table 5 shows the distribution of the four main coordinators in the two corpora. In both corpora *and* appears to be the most frequent coordinator; however, there is a difference between them with regard to the second spot. In the CEUE the adversative *or* is second, followed by *but*, while in the BNC *but* occurs about 2.5 times more frequently than in the CEUE (in bold). It is interesting to note that the conjunction *nor* is five times less frequent in the CEUE than in the BNC.

A similar examination of the subordinating conjunctions in the CEUE and the BNC Written reveals further variation between the two corpora, confirming the general impression readers of EU texts may have regarding these features of textual organization. A full list of the distribution of subordinating conjunctions in the CEUE and their number of occurrences in the BNC Written can be found in Supplementary material. The analysis showed that there is considerable difference in the use of some causal, clarifying, and additive conjunctions in the two corpora. For example, some causal conjunctions such as *with a view to*, *in order to*, *so as to*, *to this end* occur strikingly more frequently in the CEUE than in the BNC Written. Clarifying and continuative conjunctions as for example, *such as*, *in particular*, *namely*, *regarding*, *as regards* are typical of the

CEUE and not at all important in the BNC Written. By contrast, some adversative and hypothetical conjunctions, such as *although, though, if, unless*, are used much less frequently in the CEUE as compared to the BNC Written. Another interesting finding is that some conjunctions that are relatively frequent in the BNC Written are totally missing from the CEUE. These conjunctions include *in spite of, as though, by comparison, next, in conclusion*. These findings lend further empirical support to the claim that corpus-based studies are highly relevant in ESP course design, materials development and teaching, and that the careful analysis of frequency data may be used to decide which items to include and not to include in the syllabus (Biber & Reppen, 2002; Flowerdew, 1993; Mudraya, 2006; Trebits, 2008).

5.3. Data-driven learning and the discovery of usage patterns

Information on actual word use in context can be obtained with the help of concordance lists. Concordance lists are not only useful in checking frequency data, but they also provide ample instances of actual language use in context that can be directly incorporated into teaching and test materials (Flowerdew, 1993; Mudraya, 2006; Thurstun & Candlin, 1998). Data-driven learning activities can be exploited to teach many areas of ESP where learning can be greatly enhanced by using a corpus especially built to meet the specific needs of the learners. In this Section 1 will present some examples of how concordancing may help design tasks for DDL activities that can be used in the ESP classroom and/or included in [Supplementary material](#). The following activities all concentrate on the study of some of the most important conjunctions in the CEUE.

The conjunction *with a view to* appears to be extremely frequent in the CEUE as it occurred a total of 48 times, corresponding to 240 occurrences when normed to one million words. By comparison, the BNC Written contains only 8 instances of the same conjunction (also normed to one million words). This difference alone justifies its inclusion in a syllabus focusing on EU English. Selected sample lines from the concordance in [Fig. 2](#) below show the usage patterns of this conjunction.

Activity 1. This subordinating conjunction indicates a causal/purpose relationship between clauses. Study the concordance list to find out how to use it in a sentence by examining the words and phrases that follow it.

Answer Key:

1. with a view to + ing form of verb

e.g.: ... *the Council reached political agreement **with a view to adopting** the amended proposal of the European Parliament...*

2. with a view to + noun phrase

e.g.:# 1 ... *Legislative work continued **with a view to the adoption of a recommendation** of the European Parliament...*

2 ... *The summit's main achievements were the opening of negotiations **with a view to an association agreement** with Central America...*

1 ...	in the 22 official languages of the Union, with a view to facilitating in-depth discussion and debate
2 ...	an action plan on civilian aspects of the ESDP with a view to the improvement of civilian capabilities.
3 ...	urged the Commission to explore all avenues with a view to improving the (1) Directive 2006/43/EC
4 ...	Legislative work continued with a view to the adoption of a recommendation of the
5 ...	European economic and monetary union (4). With a view to the celebration of the 50th anniversary
6 ...	to evaluate the damage and requirements with a view to assessing Lebanon's current priorities
7 ...	discussion and future decisions in this area with a view to achieving sustainability, competitiveness and
8 ...	In July 2005, with a view to implementing the strategy, the Commission
9 ...	implementation of Joint Action 2002/589/CFSP with a view to a European Union contribution to combating
10 ...	Work is in progress with a view to presenting a strategic analysis of the
11 ...	into a comprehensive single regulation, with a view to streamlining and simplifying the legal
12 ...	organisation of Member States' consular offices with a view to implementing the visa information system.
13 ...	a number of issues to be examined with a view to amending the agreement on trade,
14 ...	Commission had presented on 30 June 2005 (1) with a view to creating, within the European Union,
15 ...	partnership for European tourism' (5). With a view to analysing the competitiveness of this industry
16 ...	revised the regulation establishing the Centre (1) with a view to further legislative simplification in that area
17 ...	In a resolution adopted on 15 June with a view to the 2006 Review Conference on the United
18 ...	an end to the system of fixed exchange rates. With a view to setting up their own monetary union, EU

Fig. 2. Concordance sample of *with a view to*.

Activity 2. The concordance also helps to point out the possible positions of this conjunction in a sentence. Study the list and decide whether this conjunction can be used in initial, medial or final position in a sentence. Collect examples. Which position is the most common in the CEUE?

Answer Key:

As can be seen in Fig. 2, in the vast majority of the sentences, this conjunction is in mid-sentence position. In a few other cases, however, it also appears in initial position as in the following example since the clauses of the sentence can be reversed:

... *With a view to setting up their own monetary union, EU countries decided...*

It is useful to consider the concordance lists for some other causal conjunctions that were found to be considerably more frequent in the CEUE than in the BNC Written (the occurrences below are normed to one million words for both corpora): *in order to* (325 occurrences in the CEUE and 12 in the BNC Written), *so as to* (130 occurrences in the CEUE and 23 in the BNC Written), and *to this end* (50 occurrences in the CEUE and three in the BNC Written). Figs. 3–5 present sample concordance lines of the conjunctions *in order to*, *so as to* and *to this end*. The concordance-based activities that follow help students get an insight into the syntactic and usage patterns associated with these conjunctions in the CEUE.

Activity 3. These three conjunctions express very similar relations between clauses, namely a causal/purpose relationship. Compare the three sets of data in terms of the positions of these conjunctions in the sentence.

Answer Key: *So as to* cannot appear in sentence initial position, whereas *in order to* can. *To this end* can appear in all three positions (sentence-initial, mid-sentence and sentence-final).

Activity 4. As you can see in the examples, both *in order to* and *so as to* are followed by verbs. Check the verb forms that follow the conjunctions. What verb form is used in the sentences?

Answer Key: Both conjunctions are followed by the infinitive verb form without “to”.

Activities using language data from corpora may help students gain access to the way language is really used through a large number of authentic instances of language. The data-driven approach appears to have numerous merits in teaching ESP as well as general language. Not only do students have the opportunity to use and learn the real language required in their profession, but the activities also train them to become

1	essential to a tariff union, was progressively created in order to ensure that wherever goods were imported
2	... amending the decision establishing the network in order to solve these problems
3	. goods vehicles for the use of certain infrastructures in order to reduce the differences between Member
4	framework for the programming of development aid in order to improve its effectiveness
5	in March, followed by the publication of its minutes, in order to raise the profile of the group's work and to
6	and presents those the Commission intends to adopt in order to promote the European Community's self-
7	.. support for it based on the added value it provides. In order to strengthen the Community's contribution to
8	... anywhere in the world. Resolution of disputes in order to avoid lengthy and costly court cases
9	... majority voting in the Council; on 15 September, in order to improve public access to information on its
10	examine the various monitoring methods available, in order to increase their effectiveness and improve the
11	.. agreements on their status of associate member in order to cooperate with the European GNSS
12	of measures be undertaken by the Member States in order to achieve these goals.
13	Heads of State to step up their reform efforts in order to accelerate their progress on the road to the

Fig. 3. Concordance sample of *in order to*.

1	... companies can lodge a complaint with the Court so as to have this failure to act officially recorded.
2	... people access to training and apprenticeships, so as to improve their chances of finding work ...
3	... 430 agreements have already been amended so as to recognise the principle of a Community ...
4	... common agricultural policy has been reformed so as to reward farmers for improving ...
5	... cooperate in the area of policing and justice so as to make Europe safer and more secure.
6	..with increased controls at the EU's external borders so as to combat effectively the trafficking of people ...
7	... Russian authorities to engage with Belarus so as to start a genuine process of democratisation....
8	... candidate countries receive EU aid so as to make it easier for them to catch up ...
9	... common asylum and immigration policy, so as to make it possible to completely abolish..
10	... a common framework for fighting terrorism, so as to guarantee its citizens a high level of protection...

Fig. 4. Concordance sample of *so as to*.

1 ...	which includes concrete proposals to this end in several areas.
2 ...	adopted a package of communications to this end . At the end of August, by way of ...
3 ...	paved the way for an official proposal to this end . It also set out in particular the ...
4 ...	It one another on matters of mutual interest. To this end they have agreed to establish ...
5 ...	the policy process on drugs at Union level. To this end , it launches a wide-ranging ...
6 ...	their support for the European cause: to this end the Commission is proposing ...
7 ...	the European and national authorities. To this end , they provide for special
8 ...	pillar of the European financial system. To this end it recommends overhauling ...
9	relief to the people of Lebanon and, to this end , urged Israel to remove its ..
10 ...	payment throughout all Member States (2). To this end it advocated the introduction of...

Fig. 5. Concordance sample of *to this end*.

skilled and autonomous language learners. Having to study excerpts of language and decipher the necessary information with the help of real-life language data involves the students in a process of discovery that enhances effective language learning and language use.

6. Conclusion

This paper aimed to describe the use of conjunctions in the documents of the EU with the help of corpus linguistic techniques using the Corpus of EU English. The study also intended to show how the results of such an analysis can be relevant to teaching ESP and to syllabus and materials design. Corpus-based data-driven activities can also be part of regular language classes, improving students' learning skills while providing them with meaningful and authentic language data. Based on earlier research comparing the discourse structuring patterns of native and non-native speakers of English in a variety of written and oral genres which has shown that the speakers' L1 interferes with their discourse competence in their L2, it was argued in this paper that examining textual cohesive devices, namely the use of conjunctions, is highly relevant from a language teaching standpoint. An obvious limitation of the study is the size of the corpus, since with a larger corpus more precise observations could be made regarding the various aspects of language use examined in a series of studies on EU English. Researching more aspects of language use such as various lexico-grammatical and syntactic features will complete the project which previously involved research on lexis (Trebits, 2008). Hopefully, these studies will contribute to the design and writing of practical ESP language courses focusing on the use of English in EU documents.

Acknowledgements

I would like to thank the anonymous ESP Journal reviewers and Dr. Krisztina Károly for their thorough and critical comments on earlier versions of this manuscript.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at [doi:10.1016/j.esp.2009.04.004](https://doi.org/10.1016/j.esp.2009.04.004).

References

- Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.
- Bereczky, K. (2007). Marking logical connection in presentations. *Working Papers in Language Pedagogy*, 1, 78–95.
- Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus linguistics*. Cambridge: Cambridge University Press.
- Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. Harlow: Longman.
- Biber, D., & Reppen, R. (2002). What does frequency have to do with grammar teaching? *SSLA*, 24, 199–208.
- BNC Online <www.natcorp.ox.ac.uk>.

- Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second teaching and testing. *Applied Linguistics*, 1, 1–47.
- Connor, U. (1996). *Contrastive rhetoric. Cross-cultural aspects of second language writing*. Cambridge: Cambridge University Press.
- Cullen, R., & Kuo, I.-C. V. (2007). Spoken grammar and ELT course materials: A missing link? *TESOL Quarterly*, 41(2), 361–385.
- EU ABC Online <www.euabc.com>.
- Fischer, M. (2006). *Translation(policy) and terminology in the European Union*. Paper presented at the International Conference on Terminology, Lessius University, Antwerpen.
- Flowerdew, J. (1993). Concordancing as a tool in course design. *System*, 21(2), 231–244.
- Gardner, D., & Davies, M. (2007). Pointing out frequent phrasal verbs: A corpus-based analysis. *TESOL Quarterly*, 41(2), 339–359.
- Gilmore, A. (2004). A comparison of textbook and authentic interactions. *ELT Journal*, 58(4), 363–374.
- Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English*. London: Longman.
- Heatley, A., Nation, I. S. P., & Coxhead, A. (2002). *Range and frequency programs software*. <http://www.vuw.ac.nz/lals/staff/Paul_Nation>.
- Hughes, R., & McCarthy, M. (1998). From sentence to discourse: Discourse grammar and English language teaching. *TESOL Quarterly*, 32(2), 263–287.
- Jablonkai, R. (2008). *Defining EU discourse. Steps towards the EU English discourse corpus*. Paper presented at the SEAS PhD Conference, Eötvös Loránd University, Budapest.
- Jablonkai, R., Erdei, J., Fekete, É., Homolya, K., & Nagy, É. (2006). *Angol Magyar Tematikus Gazdasági Szókincstár. [English Hungarian thematic dictionary of economics]*. Budapest: Aula Kiadó.
- Jenkins, J. (2001). 'Euro-English' accents. *English Today*, 68(17), 16–19.
- John, T. DDL page online. <web.bham.ac.uk/johnstf/timconc.htm>.
- Johns, A. M. (1980). Cohesion in written business discourse: Some contrasts. *English for Specific Purposes*, 1(1), 35–44.
- Károly, K. (2007). *Szövegten és fordítás. [Text linguistics and translation]*. Budapest: Akadémiai kiadó.
- Kennedy, G. (1998). *An introduction to corpus linguistics*. London: Longman.
- Klaudy, K., & Bart, I. (2003). *EU-fordítóiskola. Európai uniós szövegek fordítása angolról magyarra [Translating EU texts from English into Hungarian]*. Budapest: Corvina.
- Laufer, B., & Hulstijn, J. (2001). Incidental vocabulary acquisition in a second language: The construct of task-induced involvement. *Applied Linguistics*, 22(1), 1–25.
- Leech, G., Rayson, P., & Wilson, A. (2001). *Word frequencies in written and spoken English*. London: Longman.
- Liu, D. (2003). The most frequently used spoken American English idioms: A corpus analysis and its implications. *TESOL Quarterly*, 37(4), 671–700.
- McArthur, T. (1998). *The English languages*. Cambridge: Cambridge University Press.
- McArthur, T. (2003). World English, Euro-English, Nordic English? *English Today*, 73(19), 54–58.
- Mishan, F. (2004). Authenticating corpora for language learning: A problem and its resolution. *ELT Journal*, 58(3), 219–227.
- Modiano, M. (2001). A new variety of English. *English Today*, 68(17), 13–14.
- 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(2), 235–256.
- Nickerson, C., Gerritsen, M., & van Meurs, F. (2005). Raising student awareness of the use of English for specific business purposes in the European context: A staff–student project. *English for Specific Purposes*, 24(3), 333–345.
- Pym, A. (2000). The European Union and its future languages: Questions for language policies and translation theories. *Across Languages and Cultures*, 1(1), 1–17.
- Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1985). *A comprehensive grammar of the English language*. New York: Longman.
- Rogerson-Revell, P. (2006). Using English for international business: A European case study. *English for Specific Purposes*, 26(1), 103–120.
- Scott, M. (1996). *Wordsmith lexical analysis tools software*. Oxford: Springer, <<http://www.lexically.net/wordsmith/>>.
- Seidlhofer, B. (2001). Towards making 'Euro-English' a linguistic reality. *English Today*, 68(17), 14–16.
- Seidlhofer, B. (2005). English as a lingua franca. *ELT Journal*, 59, 339–341.
- Seidlhofer, B. (2007). Common property: English as a lingua franca in Europe. In J. Cummins & C. Davison (Eds.), *International handbook of English language teaching* (pp. 137–153). New York: Springer.
- Swales, J. M. (1990). *Genre analysis: English in academic and research settings*. Cambridge: Cambridge University Press.
- Thurstun, J., & Candlin, C. N. (1998). Concordancing and the teaching of the vocabulary of academic English. *English for Specific Purposes*, 17(3), 267–280.
- Trebits, A. (2008). English lexis in the documents of the EU – A corpus-based exploratory study. Eötvös Loránd University: *Working Papers in Language Pedagogy*, 2. <<http://lagped.elte.hu/Wopalpindex.htm>>.
- Trosborg, A. (1997). Translating hybrid political texts. In A. Trosborg (Ed.), *Text typology and translation* (pp. 145–159). Amsterdam: John Benjamins.
- Truchot, C. (2002). *Key aspects of the use of English in Europe*. Strasbourg: Council of Europe.
- Vergaro, C. (2004). Discourse strategies of Italian and English sales promotion letters. *English for Specific Purposes*, 23(2), 181–207.

Anna Trebits is a PhD student at Eötvös University, Budapest, Hungary. She holds an MA in English and French Language, Literature and TESOL. She has been teaching English to young adults and adults for 10 years in Hungary, Portugal and Luxembourg. Previously she worked for an EU institution in Luxembourg.