TYPICAL COLLOCATIONS IN THE JUDGMENTS ON APPEAL OF THE EUROPEAN COURT OF JUSTICE

Daiva Macko

Statistics Lithuania, Gedimino pr. 29, LT-01500 Vilnius, Lithuania
E-mail: *daiva.macko@stat.gov.lt*

The research paper aims to identify typical collocations frequently used in the appellate judgments of the European Court of Justice (ECJ) and to compare their use with the general English language. The methodological guidelines of corpus linguistics were followed in the course of investigation. The research focuses on the analysis of the right verbal collocates of *Court.* The British National Corpora (BNC) was used as the source of general English. The quantitative methods include applying statistical measures (MI score and log-likelihood) to test the significance of the extracted collocations and to compare the calculated values with those of corresponding collocations in the BNC. The qualitative part of the research focuses on classifying the most frequent collocations of the chosen syntactic pattern (NOUN + VERB) into structural (grammatical) and semantic patterns. The results show that typical collocations used in the appellate judgments of the ECJ differ from the general English language in terms of frequency and statistical significance and exhibit unique semantic characteristics, therefore suggesting that there are considerable lexical differences between legal and general English that should be taken into account in teaching and learning.

Keywords: collocation, colligation, corpus, semantic patterns, statistical significance.

doi:10.3846/cpe.2011.08

Aim

The present research aims to select the most frequent collocations used in a chosen field of specialist language, as well as to compare the results obtained with respective data of general language. It is also sought to discuss structural (grammatical) and semantic properties of the typical collocations extracted in order to find out if they exhibit any genre-specific features. The differences in the use of selected collocations between the general and specialised English are supposed to illustrate that specific collocational competence should be involved in teaching and learning specialised English in general and legal English in particular.

Previous research

It is generally agreed that the origins of the concept of collocation in linguistics lie in Firth's definition of the phenomenon as 'actual words in habitual company' (Firth 1957: 14 quoted in Kennedy 1998: 108), or 'the company words keep' (Firth quoted in Hill 2000: 48). In the current research, a statistical approach to collocation is followed rather than a semantically-based approach. A statistically-based concept of collocation relies on the application of computational tools to large corpora and extraction of recurrent patterns of words (Siepmann 2005: 410–411). The statistical approach was advocated and developed by Sinclair (Crowther *et al*. 2002:

58). The frequency criterion seems to be acceptable to many linguists and thus can be stated to lie at the heart of the statistically-based concept of collocation (see Bartsch 2004: 59–60; Otani 2005: 5; Hanks 2008: 222; Lewis 2000: 127, etc.). For example, Biber *et al*. define collocations as statistical associations of words that often co-occur together (Biber *et al.* 1999: 988). In principle, the statistical approach to collocation implies that the validity of results obtained is directly dependent upon the number of recurrent patterns in a large corpus. In other words, the greater the co-occurrence of certain words in the same corpus, the more likely they will collocate with each other than appear separately. Although it sounds reasonable, this point is criticised by Siepmann (2005: 411), who notices that it remains unclear at which point frequency becomes significant enough. As a result, it is often considered that there are no clear boundaries to mark the significance of collocates (Otani 2005: 5; Kennedy 1998: 117).

As regards the form and content of collocation (i.e. the number and nature of elements that constitute it), collocations composed of the so-called content words are generally referred to as lexical (Wei 1999: 8; Lewis 2000: 134), distinct from grammatical collocations involving a grammatical structure or containing prepositions. The latter are usually referred to as colligation (see Siepmann 2005: 411–419; Sinclair 2000: 200; Hanks 2008: 222; Hoey and Brook 2008: 294; Hoey 2000: 234). In theory, lexical collocations are considered to combine two equal lexical components; yet these combinations of words are almost always embedded in certain grammatical structures, thus the number of the constituents of collocation is actually more than two items (Lewis 2000: 134).

The predominant semantic properties of collocation stem from its contextual origins and the importance of repetition in a text. It is considered that through constant repetition and repeated co-occurrences textual and intertextual meaning is formed (Siepmann 2005: 409; Stubbs 2001: 157). It is worth remembering that Firth also undermined the repetition in language as a source of typicality ('typical, recurrent and repeatedly observable', Firth 1957: 35 quoted in Tognini-Bonelli 2001: 164). According to Tognini-Bonelli, the very concept of collocation arises from the above mentioned theoretical premises. She also notes that the Firthian theory postulates the priority of lexis over grammar (collocation should be observed first and colligation inferred after, as it is a more abstract feature), which obviously has implications for language teaching by shifting the focus to lexis rather than grammar[1].

In the field of collocational studies, corpus linguistics is chosen by many researchers as a methodological guideline. Probably, this choice is largely motivated by the importance of frequency criterion. Unexpectedly, the findings in this area have revealed that a number of collocations extracted from corpora are not accessible to intuition, i.e. the users of language are to some extent unaware of their own collocational competence and the patterns that they produce (Widdowson 2000: 6). Therefore it is all the more important to study collocational patterning in order to improve the teaching and learning processes of a specialist vocabulary. In addition, these findings distinguish corpus linguistics as a valuable quantitative method; the observations made on the basis of this method are empirical by nature and provide the results with an objective quality. Sinclair has often emphasized the importance of objective observance of language in use in order to find 'evidence', or facts about language and its regularities (for example, see Sinclair 1991: 39).

There are various kinds of corpora distinguished, but the most relevant distinction in this research is that between general and specialized corpora. The former contain texts from different genres and often include spoken and written language, while specialized corpora are designed for specific research and are confined to language used only in particular kinds

[1] *For research based on the so-called lexical approach see, e.g., Lewis, M. 2000. Teaching Collocation. England: Language Teaching Publications.*

of texts or situations (Kennedy 1998: 19–20; Paltridge 2006: 156–157).

As a result of investigation in the field, a fairly novel concept of 'collocational competence' has emerged. It is often epmphasized as a vital skill for adequate knowledge of language (see, for example, Hill 2000: 49; Juknevičienė 2008: 119). Consequently, mastery of special languages can also be regarded as largely dependent on collocational competence, since lexis is an important attribute of a genre. Therefore, a lot of research focuses on typical collocations in special languages as opposed to common collocations in general language.

However, it seems that little research has been done on typical collocations in the legal language in general and in the legal language of the EU[2]. As regards the legal English of the Commonwealth countries, Bhatia (1993) and Maley (1994) have discussed its genre-specific traits in detail. Bhatia defines judgments as one of the most conventionally standardized disciplinary genres (together with legislation and case-law) in law (Bhatia 2000: 82). Nevertheless, Bhatia notes that legal English in Europe differs considerably from the countries of common law system (Bhatia 1993: 139), while Maley also emphasizes that particularly at the appellate level structural differences of continental and common law systems are evident (Maley 1994: 44). The nature of the EU law system is distinct both from continental and common law. Yet, to the best of my knowledge, so far the discoursal and linguistic characteristics of the EU legal English have not been systematized yet. Collocational studies in this field could provide valuable insights into its lexical characteristics.

Data and methods

The research is based on the analysis of a corpus composed of 50 judgments on appeal of the European Court of Justice. The size of the corpus is 528,073 words. The judgments are available on the website <http://eur-lex.europa.eu> in the Internet. The chosen judgments were retrieved from this database in the following sequence: filtering the data by specifying a file category (case-law); narrowing the selection to the documents issued by the Court of Justice; and filtering files according to the type of procedure – choosing appeal procedure. As a result of this search, 432 judgments were available at the moment of selection (3 November 2009). The time span of these judgments begins from 1 October 1991; the database is frequently updated. In order to compare the results with the data from a larger, general corpus, the British National Corpus (BNC, 100 million words), available at <http://corpus.byu.edu/bnc>, was used.

The corpus was composed with the aim to analyse the most recent available data, as it was expected that this material would be the most representative of the current use of the legal English language of the European Union (EU) institutions (representativeness is commonly distinguished as one of the key characteristics of a corpus, see Tognini-Bonelli 2001: 52–62; McEnery and Wilson 2001: 30). Therefore, the most recent judgments were chosen, covering the time span from 21 February 2008 until 10 September 2009. The authorship of the selected texts is attributable to groups of persons rather than a single author, because a judgment is arrived at by a Chamber composed of several judges.

The methods used were both quantitative and qualitative. The programme WordSmith Tools (WS), Version 5.0 was used to extract collocations and calculate their frequency and statistical significance scores. Computational tools available on the Internet were also applied. The qualitative part of the research was combined with computational analysis and involved manual scrutiny of relevant (i.e. statistically significant and most frequent) collocations. The focus was centred upon the classification of data into grammatical and semantic patterns.

[2] *Although there is literature available on general lexical and syntactic properties of the language of the law (see Vystrčilova (2000); Gibbons (1994); Cao (2007); Ingels (2006)).*

Results and discussion

Quantitative analysis. Initially, quantitative research was carried out in order to select the material for qualitative analysis. The concepts underlying this part of the research are the following.

Concordance – 'a comprehensive listing of a given item in a corpus (most often a word or a phrase), also showing its immediate context' (McEnery & Wilson 2001: 197). Technically, it can also be defined as 'a list of all the words, or a certain word, used in a text or in a body of texts, together with a context in which the words appear. This context is usually no more than 7 or 8 words to the left and right of the node word' (Concordancing glossary at <http://www.nsknet.or.jp/~peterr-s/index.html>). The above mentioned context is usually referred to as a 'span' (ibid). Sinclair refers to concordance as a 'first stage in examination of an item as a node' (Sinclair *et al.* 2004: 71). Consider the following example of a machine generated concordance.

Node-word – 'the word that appears in the middle of the screen in a list of concordances' (Concordancing glossary). Sinclair refers to a node as 'the word that is being studied' or 'a central word' in a 'machine-generated concordance' (Sinclair 1991: 105, 116). In the example below the node-word is *Court*.

Collocate – 'any word that occurs in the specified environment of a node' (Sinclair 1991: 115). The word *held* in the example above stands for a collocate. The Concordancing glossary provides a common definition based on the frequency criterion: 'a word that appears with another word more often than simple chance would suggest'.

Computational tools. The programme WordSmith Tools (WS), Version 5.0 was used to extract collocates and generate concordances and wordlists from the corpus. First, a wordlist[3] was generated in order to find out the most frequent content words in the corpus. It turned out that *Court* was the most frequent lexical item used (5386 occurrences), therefore it was chosen as a node-word. Afterwards, the concordance of the chosen node-word was generated in order to find out its collocates. Following Sinclair's recommendations (see Sinclair 1991: 106), the WS programme was set to count collocates within a span of ten words, i.e. five words to the left and five words to the right of the node-word.

[3] *A wordlist is a list of words automatically generated in both alphabetical and frequency order. (Scott 2009. WordSmith Tools Version 5.0. Liverpool: Lexical Analysis Software).*

N Concordance

24	been annexed to that statement, the **Court held** that ADM could not properly
25	279 of the judgment under appeal, the **Court held** that Schneider could not
26	on Adams v Commission, in which the **Court held** that the expiry of a limitation
27	A. In paragraph 57 of that judgment, the **Court held** that, in this case, it had not
28	56 In reaching this conclusion, the **Court held** in particular, in paragraphs
29	the application for annulment 18 The **Court held** that there was no need to rule
30	nor the exercise of those rights. The **Court held** that, contrary to the
31	Pak v Commission, paragraph 41, the **Court held**, first, that prices below
32	by the applicants. 55 Finally, the **Court held** that the damage incurred by

In order to assess the validity of the collocations extracted and to compare them with the data from the BNC, certain quantitative measures were applied. The log-likelihood calculator (available at <http://ucrel.lancs.ac.ukllwizard.html>) was used to compare the relative frequencies between the two corpora. In Table 1 below, a '+' sign indicates that the frequency of a certain collocation in the corpus outnumbers the frequency of the same collocation in the BNC. The higher the value, the more significant is the difference between two frequency scores and the lower the probability that the statistical difference is accidental.

Table 1. *Right verbal collocates of Court p < 0.0001; critical value = 15.13*

Right verbal collocates of Court	Occurrences in the chosen judgments	Occurrences in the BNC	Log likelihood
HELD	326	494	+2325.43 (overuse)
SET	119	46	+1054.44
FOUND	116	150	+854.94
FIND	6	65	+22.54
FINDS	11	18	+77.17
STATED	112	39	+1003.66
STATES	8	1	+77.71
ERRED	108	8	+1075.64
REJECTED	85	79	+666.02
REJECT	7		
DISMISSED	61	48	+491.33
DISMISS	48	10	+450.68
INFRINGED	59		
FAILED	58	23	+512.46
CONCLUDED	49	34	+402.42
CONCLUDE	9	13	+64.85
CONSIDERED	46	66	+331.93
CONSIDERS	18	33	+123.09
POINTED OUT	46		
DISTORTED	41		
NOTED	37	9	+343.04
NOTES	8	1	+77.71
COMMITTED	37	17	+321.33
APPLIED	36	15	+316.29
ACCEPTED	32	62	+216.02
REFERRED	30	31	+230.72
REFERS	5	7	+36.26
EXAMINED	30	7	+279.12
RULED	28	267	+111.63
RULE	15	29	+101.31
OBSERVED	26	10	+230.51
OBSERVES	5		
RELIED	26	10	+230.51

Continued Table 1

ANNULLED	22	3	+212.64
DECIDED	20	132	+92.98
BASED	18	21	+135.35
GIVE	18	104	+87.96
GIVES	11	23	+72.91
ORDER	17	148	+70.57
ORDERED	6	117	+16.27
EXERCISE	16	41	+100.73
ADDED	14	11	+112.79
MAKE	14	226	+46.62
RECALLED	12	2	+114.51
ASSESSED	12	1	+118.93
ASSESS	7	14	+46.90
INFERRED	12		
DECLARE	11	16	+79.15
DECLARED	8	24	+48.25
INTERPRETED	11	3	+101.96
MISCONSTRUED	10		
RECOGNISED	8	16	+53.60
DETERMINE	8	36	+42.64
DETERMINED	5	8	+35.25
MISINTERPRETED	8		
SUBSTITUTED	8	3	+71.12
DISREGARDED	8	2	+74.00
ACKNOWLEDGED	7	5	+57.24
MISAPPLIED	6		
GRANTED	6	55	+24.35
UPHELD	6	96	+18.36
ADJUDICATE	6	4	+49.57
IMPOSED	5	20	+27.68
QUASH	5	13	+31.36

In comparison with the BNC, the numerical log-likelihood values obtained show that the use of the selected collocations in my corpus is much more frequent than in the BNC, i.e. the frequency of the collocations selected is significantly higher in my corpus than in the BNC. Consequently, the results suggest that typical collocations extracted serve as generic markers, i.e. they distinguish this legal subgenre from the general language. It is worth noting that the lowest score is higher than 15.13, which means that the chances of unreliability of the calculations amount to only 0.01 percent.

In addition, the statistical MI test was applied to the selected collocations. The importance of this test is reinforced by theoretical doubt that the frequency alone does not necessarily constitute a collocation and the co-occurrence might be accidental. Thus, MI (mutual information) score is a test designed to measure the statistical significance of collocation. It 'compares the probability that the two items occur together as a joint event <…> with the probability that they occur individually', i.e. by chance (McEnery and Wilson 2001: 86). The higher the MI score, the more significant the

collocation is, whereas the values below zero show that words co-occurred by chance (ibid).

The MI test was applied to the selected data using the WordSmith Tools programme, while the BNC provides an option of displaying MI scores together with the collocations extracted. To assure that results are reliable, some linguists[4] recommend to set the cut off point for MI values at 3 and exclude the values below this point. This recommendation was followed. The results are presented in Table 2 together with the BNC MI values. The collocations selected from the analysed judgments proved to be statistically significant, while their significance in the BNC was not as high (except for *ruled* and *upheld*, which had higher MI scores). In fact, a large part of collocations in the BNC had MI values lower than 3 and thus are not included in Table 2.

[4] *See Martin Weisser's website at <http://ell.phil.tu-chemnitz.de/analysis/collocations.html>.*

Table 2. *MI score*

Node-word	Collocate	Joint frequency	MI	Joint frequency in the BNC	MI in the BNC
COURT	QUASH	5	6.62	13	6.58
	ERRED	108	6.39	8	5.57
	MISINTERPRETED	8	6.29		
	DISMISS	48	5.95	10	3.07
	MISCONSTRUED	10	5.85		
	RULED	28	5.69	267	6.24
	HELD	326	5.69	494	3.63
	MISAPPLIED	6	5.62		
	SUBSTITUTED	8	5.37		
	FOUND	116	5.35		
	INFRINGED	59	5.24		
	DISTORTED	41	5.20		
	COMMITTED	37	5.17		
	ADDED	14	5.14		
	DECLARE	11	5.07	16	3.53
	FAILED	58	5.03		
	STATED	112	5.03		
	INFERRED	12	4.95		
	DISREGARDED	8	4.91		
	POINTED	53	4.89		
	TOOK	42	4.89		
	ACKNOWLEDGED	7	4.84		
	CONCLUDED	49	4.83		
	ANNULLED	22	4.79	3	4.40
	DISMISSED	61	4.70	48	3.47
	EXPLAINED	10	4.69		
	RECALLED	12	4.65		
	ACCEPTED	32	4.65		

Continued Table 2

	REJECTED	85	4.53	79	3.76
	NOTES	8	4.53		
	NOTED	37	4.50		
	CONSIDERS	18	4.43		
	DECIDED	20	4.32		
	OBSERVED	26	4.30		
	OBSERVES	5	4.29		
	EXAMINED	30	4.26		
	SET	119	4.24		
	CONSIDERED	46	4.21		
	RECOGNISED	8	4.16		
	FINDS	11	4.01		
	CONCLUDE	9	3.98		
	INTERPRETED	11	3.75		
	ASSESSED	12	3.66		
	RELIED	26	3.54		
	REFERRED	30	3.34		
	APPLIED	36	3.25		
	UPHELD	6	3.09	96	6.79

The scope of **qualitative analysis** of the current research was restricted to a single syntactic pattern – NOUN + VERB. As the collocates analysed occurred at the right of the node-word, *Court* takes the syntactic role of a subject in such instances, while verbs following after are used in the active voice. The focus was on lexical verbal collocates, which excludes the so-called form-words from the scope of the research. Thus such auxiliaries and modals as *is, was, did, should* were ignored, just as the so-called delexicalised verbs (see Juknevičienė 2008: 120), e.g. *have, take, make, give*, etc., unless they were used in a uniform sense (e.g. *to give a judgment; to make a decision*). With the purpose of restricting the number of analysed instances, only collocates that occur not less than 5 times within the same grammatical pattern were selected.

Colligation. After scrutinizing the concordance it turned out that the collocates extracted are embedded in identifiable colligational patterns. In this paper colligations are viewed as lexico-syntactic structures, in line with Biber's definition of the phenomenon as 'lexico-grammatical associations' (Biber *et al.* 1999: 989).

According to surrounding clause elements, the extracted collocates could be grouped into two major grammatical categories: verbs taking that-complement clauses or transitive verbs. The latter can be divided into three subgroups: verbs taking a direct object; verbs taking a prepositional object and phrasal verbs taking a direct object (see Table 3; numbers in parenthesis indicate the number of co-occurrences). As regards the group of phrasal verbs, the collocations *Court took sth. into consideration* and *Court took sth. into account* can be considered as multi-word verb constructions that have idiomatic status (Biber *et al.* 1999: 427). A distinct lexico-syntactic pattern was verb + to-infinitive, manifest in a single collocation *Court failed [to do sth]*.

More often than not, the colligational patterns were rather homogenous, i.e. a verbal collocate was used in the same lexico-syntactic

pattern throughout the corpus. This was very common for verbs followed by that-clauses. Yet there were some deviations from this tendency: *Court accepted* either introduced a that-clause or an object of the verbal collocate, while verbs *hold* and *uphold,* although semantically related, were subject to different lexico-syntactic patterns.

In some cases, even the lexical content to the right of the collocation was rather uniform, suggesting patterns for trinomial rather than binomial collocations, for example, *Court erred in law* (85) or *Court should dismiss an appeal* (36). These instances match Siepmann's remark that, even though there is a tendency to view collocation as a binary unit, some findings suggest that quite often a trinomial structure prevails which cannot be reduced to a binary one (see Siepmann 2005: 412–417).

Table 3. *The colligational patterns of right verbal collocates of Court*

Verbal collocates + that-clauses	Verbal collocates + direct object	Verbal collocates + prepositional object	Phrasal verbal collocates + direct object	Verbal collocates + to-infinitive
Court held *that (320)* Court found *that (116)* Court finds *that (10)* Court stated *that (112)* Court states *that (7)* Court concluded *that (49)* Court considered *that (46)* Court pointed out *that (39)* Court noted *that (37)* Court observed *that (17)* Court decided *that (15)* Court took the view *that (14)* Court added *that (12)* Court accepted *that (12)* Court ruled *that (11)* Court inferred from sth. *that (11)* Court acknowledged *that (7)* Court recalled *that (6)* Court … declare *that (5)* Court explained *that (5)* Court recognised *that (5)*	Court rejected *sth. (85)* Court dismissed *sth. (61)* Court … dismiss *sth. (48)* Court infringed *sth.(59)* Court distorted *sth.(41)* Court committed *sth. (37)* Court applied *sth. (35)* Court examined *sth. (23)* Court annulled *sth. (22)* Court accepted *sth. (17)* Court … give *sth. (17)* Court gives *sth. (11)* Court … exercise *sth. (16)* Court … make *sth. (14)* Court … order *sth. (13)* Court assessed *sth. (12)* Court interpreted *sth. (11)* Court misconstrued *sth. (10)* Court misinterpreted *sth. (8)* Court disregarded *sth. (8)* Court … determine *sth. (7)* Court misapplied *sth. (6)* Court upheld *sth. (6)* Court granted *sth. (6)* Court imposed *sth. (5)* Court … quash *sth. (5)*	Court erred in sth. *(105)* Court referred to sth. *(27)* Court relied on sth. *(27)* Court based sth. on sth. *(13)* Court … rule on sth. *(12)* Court ruled on sth. *(11)* Court substituted sth. for sth. *(8)* Court took account of sth. *(8)*	Court set aside sth. *(95)* Court set out sth. *(14)* Court took sth. into consideration *(8)* Court took sth. into account *(6)*	Court failed *to do sth. (58)*

Semantic patterns. In addition to lexico-syntactic patterns, the collocations analysed were also subject to certain patterns of attitudinal meaning. In linguistics, communication of 'feelings, attitudes, value judgments, or assessments' are termed stance (Biber *et al.* 1999: 966). While the expression of feelings, naturally, is not welcome in juridical settings, the expression of attitudes and assessments seems to be involved in the argumentation of the parties to the proceedings.

The genre of appellate judgments per se presupposes a negative evaluation of the court of first instance's decisions and consists of numerous indications of errors in the judicial reasoning. This negative stance is expressed in verbs that have an element of negative evaluation, for example, *to* ***err****, to* ***infringe****, to* ***distort****, to* ***misconstrue***, etc. Consider the list of collocations marked for negative stance within the most typical context of the corpus:

- ***Court erred in sth.*** *(105):* ***Court erred*** *in law (85)*
- ***Court infringed sth.*** *(59):* ***Court infringed*** *an Article / an obligation / a principle / a provision, etc.*
- ***Court failed to do sth.*** *(58):* ***Court failed*** *to explain / to respond to arguments / to state adequate reasons for finding, etc.*
- ***Court distorted sth.*** *(41):* ***Court distorted*** *the facts/the evidence, etc.*
- ***Court misconstrued sth.*** *(10):* ***Court misconstrued*** *the legal criteria / the provision, etc.*
- ***Court misinterpreted sth.*** *(8):* ***Court misinterpreted*** *an Article / a concept, etc.*
- ***Court misapplied sth.*** *(6):* ***Court misapplied*** *the article / the provisions / the test, etc.*
- ***Court disregarded sth.*** *(8):* ***Court disregarded*** *the fact / the Article, etc.*

Such collocations were always followed by evaluative phrases in the analysed judgments. For example:

1. By the second part of this ground of appeal, the ***appellant claims that*** *the* ***Court*** *of First Instance* ***failed to*** *state adequate reasons for rejecting the appellant's arguments <…> ;*

2. … the ***appellants submit that*** *the* ***Court*** *of First Instance* ***misconstrued*** *the burden and standard of proof <…> ;*

3. Lastly, the ***appellant alleges that*** *the* ***Court*** *of First Instance* ***failed to*** *have regard to the fact that <…> ;*

4. The ***appellant claims that****, by declaring inadmissible certain documents produced for the first time before the Court of First Instance, that* ***court infringed*** *Article 63 of Regulation No. <…> ;*

5. In the second place, the ***appellant submits that*** *the* ***Court erred in law*** *in misconceiving the consequences of invalidity;*

6. <…> the ***appellant claims that*** *the* ***Court*** *of First Instance* ***distorted*** *the evidence on which it based its analysis <…> ;*

7. By its first plea, the ***appellant claims that*** *the* ***Court*** *of First Instance* ***misconstrued*** *the concept of 'internal competition' <…> .*

Interestingly, after examining the linguistic context of such collocates, it turned out that the appellant's arguments account for a large portion of the right verbal collocates of *Court* marked for negative stance (e.g. consider the above given instances). The regularities observed stem from the peculiarities of juridical settings. At the appellate instance, the appellant is to point out his reasons for appealing. Specifically, he is supposed to give legitimate reasons for his dissatisfaction with the court's decision. Therefore, he has to continuously refer to the court of first instance's arguments or actions that he considers to be erroneous in some respects. Naturally, the appellant's argumentation involves evaluative aspects, thus accounting for the verbal collocates of *Court* marked for negative stance, while linguistically, that-clauses serve as a convenient syntactic pattern to structure and present the appellant's claims. The figure provides a schematic illustration of the above discussed observations.

Before proceeding to the discussion of the remaining collocates, a fairly novel concept of semantic prosody needs to be considered. In

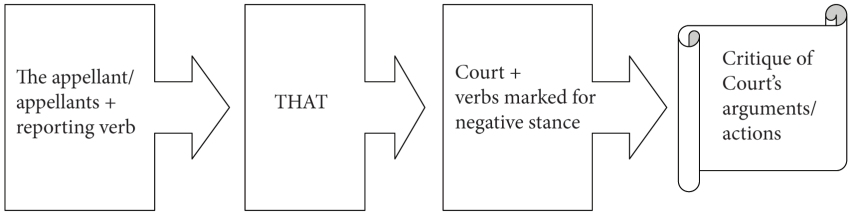

Fig. 1. *The structure of appellant's arguments*

the field of collocational studies this linguistic term is understood as 'an attitudinal and pragmatic meaning' opposed to referential meaning (Sinclair 2000: 200). Jackson understands it as 'particular negative or positive connotations' (Jackson 2002: 16). It seems that this novel term has a particular purpose in studying collocation and relates to the special peculiarity of words to collocate seemingly unpredictably and also for one collocate to enhance certain semantic aspects of the other collocate. Sometimes, these aspects can be unexpected, thus referred to as 'latent categories of meaning' by Sinclair (2000: 198). For example, Hoey states that the verb *happen* is more likely to associate with unpleasant events (2000: 232).

In the instances discussed above, right verbal collocates of *Court* convey negative stance explicitly. The negativity is either encoded morphologically (in the prefixes *mis-*, *dis-* in words *misconstrued, misinterpreted, misapplied, disregarded*) or lexically (*erred, failed, infringed, distorted*). Yet another group of collocates could be distinguished consisting of verbs that do not denote negative actions per se, yet they occur solely in the semantic environment of evaluation or of claiming something to be erroneous or illegal, for example, ***Court committed*** *an error of law / a manifest error* (this phrasing is manifest in as many as 36 instances out of 37 co-occurrences). The latter collocations could be regarded as having a negative semantic prosody. Similarly, the collocation ***Court applied***, although itself rather neutral, can be said to have an evaluative semantic prosody, since it is used in the statements of evaluative nature: it is frequently surrounded by words *correctly / incorrectly* and similar expressions. Consider the examples of negative and evaluative semantic prosody:

1. Salzgitter submits that the ***Court*** *of First Instance* ***committed an error of law*** *in finding that <…> ;*

2. The Commission also submits that the ***Court*** *of First Instance* ***committed an error of law*** *in accepting that <…> ;*

3. Moreover, in its examination of the contested decision, the ***Court*** *of First Instance itself* ***committed manifest errors*** *of assessment and fundamentally misconstrued the evidence <…> ;*

4. Finally, as regards the third part, the appellant maintains that the ***Court*** *of First Instance* ***incorrectly applied*** *the principles identified in the Baby-Dry judgment;*

5. In the alternative, Impala submits that the ***Court*** *of First Instance* ***applied*** *the* ***correct*** *test for establishing market transparency in the judgment under appeal, namely <…> ;*

6. The Commission submits that the ***Court*** *of First Instance* ***erroneously applied*** *the case-law mentioned in paragraph 22 <…> ;*

7. In this ground of appeal, the Kingdom of Belgium alleges that the ***Court*** *of First Instance* ***wrongly applied*** *the principle of proportionality in considering that <…> .*

Interestingly, statistical measures applied showed that the collocates marked for negative stance proved to be the most significant (see MI score results in Table 2). Having in mind

the Firthian theory (repeated co-occurrences in language are a source of typicality, see Tognini-Bonelli 2001: 164), these collocates can be regarded as typical to the sub-genre under investigation, i.e. they can be viewed as the most characteristic lexical elements of the appellate judgments of the ECJ. As such, they are specific and therefore do not occur in general English as frequently; in fact, most of them (*infringed; distorted; misconstrued; misinterpreted; misapplied*) were not available in the BNC at all. These findings suggest that the narrower the genre, the more specific collocations it tends to exhibit in relation to general language.

Other collocations, in contrast, seem rather neutral in terms of stance and are used to express the reasoning of the parties to the proceedings. The latter verbal collocates resemble what Biber et al. classify as mental and communication verbs (Biber *et al.* 1999: 362–363). Just as Biber *et al.* suggest (ibid: 660–670), mental and communication verbal collocates were frequently subject to colligational patterns with that-clauses which introduced reported statements (e.g. ***Court stated*** *that,* ***Court considered*** *that,* ***Court concluded*** *that,* ***Court noted*** *that* etc.).

Some verbal collocates indicate the court's actions rather than its reasoning and argumentation processes; they are used to name certain procedural steps in decision-making process and relate to the court as a procedural body, for example: ***Court dismissed*** *an appeal*, ***Court ruled*** *on a plea*, ***Court examined*** *a dispute* etc. Following Biber's classification, these verbal collocates would fall under the heading of activity verbs (ibid: 362). Collocations with these verbs can also be regarded as neutral in terms of stance.

Conclusions and implications for further research

The following conclusions can be drawn from the research carried out:

1. The current study showed that the language of the judgments on appeal of the European Court of Justice is significantly different from the general English language:
   - The results of statistical analysis prove that in terms of frequency and statistical significance of the analysed collocations the language of the appellate judgments in the EU law is remarkably different from general English.
   - The qualitative analysis revealed that collocations analysed exhibit grammatical and semantic properties that allow classifying them into colligational and attitudinal patterns.
2. The semantic analysis suggests that the genre of appellate judgments is unique because:
   - it provides collocations that express numerous ways of saying that the court was wrong;
   - the results obtained from statistical analysis show that the right verbal collocates of *Court* marked for negative stance accounted for the most significant collocations throughout the corpus and in comparison to the BNC, thus proving that collocations serve as a source of typicality and suggesting that perhaps some ways of expressing the wrongfulness of court's actions is specific to the English of the EU law only.

The current research suggests several implications:

1. The results suggest that typical collocations extracted serve as generic markers, i.e. they distinguish the legal subgenre of appellate judgments of the ECJ from the general language.
2. The current research reaffirms that the study of legal English should be specialised, i.e. it should differ from teaching general English language.
3. The results suggest that a specific collocational competence should be involved in producing the language of the EU law.

Since the research was limited in various respects, suggestions for further research arise:

1. The size of the corpus could be expanded in order to achieve more valid results, as it is supposed that larger corpora provide rarer uses (see McEnery & Wilson 2001: 80). For example, the judgments could cover a larger time span.
2. Having in mind the variety of subgenres of legal discourse, the appellate judgments

analysed cover a relatively narrow area of the language of the EU law. In comparison, different subgenres could be considered.

3. A parallel Lithuanian corpus could be composed for further contrastive studies of the genre of appellate judgments of the ECJ, since all of the selected judgments are translated into Lithuanian.

4. It would be interesting to investigate the English language of appellate judgments in the countries of common law in order to compare it with the legal English of the EU law.

References

Bartsch, S. 2004. *Structural and Functional Properties of Collocation in English* [Accessed on 17 August 2009]. Available from Internet: <http://books.google.com/>.

Biber, D. S.; Johansson, G.; Leech, S.; Finegan, C. and E. 1999. *Longman Grammar of Spoken and Written English*. Harlow: Pearson Education Limited.

Bhatia, V. K. 1993."Legal Duscourse in Professional Settings", in *Language Use in Professional Settings.* London: Longman, 101–144.

Bhatia, V. K. 2000. "Discourses of Business and Law: is there a Common Core?", in Dvorakova, J. (Ed.). *Proceedings of the LSP forum '99, Prague, Czech Republic, September 17–19, 1999*. Prague: Asocisace učitelu angličitiny, 81–86.

Cao, D. 2007."Translating Law (Topics in Translation)", in *Multilingual Matters LTD*, 20–23.

Concordancing glossary [Accessed on 6 March 2010]. Available from Internet: <http://www.ns-knet.or.jp/~peterr-s/index.html>.

Crowther, J.; Dignen, S. & Lea, D. 2002. "Oxford Collocations Dictionary for Students of English", *International Journal of Lexicography* 16(1): 57–61.

Gibbons, J. 1994. "Language Constructing Law", in Gibbons, J. (Ed.). *Language and the Law*. London and New York: Longman, 3–10.

Hanks, P. 2008. "The Lexicographical Legacy of John Sinclair", *International Journal of Lexicography* 21(3): 219–229. doi:10.1093/ijl/ecn031

Hill, J. 2000. "Revising priorities: from Grammatical Failure to Collocational Success", in Lewis, M. (Ed.). *Teaching Collocation*. England: Language Teaching Publications, 47–68.

Hoey, M. 2000. "A World Beyond Collocation: New Perspectives on Vocabulary Teaching", in Lewis, M. (Ed.). *Teaching Collocation*. England: Language Teaching Publications, 224–243.

Hoey, M. and Brook, M. 2008. "Lexicography, Grammar and Textual Position", *International Journal of Lexicography* 21(3): 293–309. doi:10.1093/ijl/ecn025

Ingels, M. B. 2006. "Specific Characteristics of Legal English Writing", in *Legal English Communication Skills*. Acco Leuven/Veorburg, 5463.

Jackson, H. 2002. *Lexicography. An Introduction*. London and New York: Routledge. doi:10.4324/9780203467282

Juknevičienė, R. 2008. "Collocations with High-frequency Verbs in Learner English: Lithuanian Learners vs Native Speakers", *Kalbotyra* 59(3): 119-126.

Kennedy, G. 1998. *An Introduction to Corpus Linguistics.* London and New York: Longman.

Lewis, M. 2000. "Language in the Lexical Approach", in Lewis, M. (Ed.). *Teaching Collocation*. England: Language Teaching Publications, 126–154.

Maley, Y. 1994. "The Language of the Law", in Gibbons, J. (Ed.). *Language and the Law*. London and New York: Longman, 11–49.

McEnery, T. and Wilson, A. 2001. *Corpus Linguistics.* Edinburgh: Edinburgh University Press.

Otani, H. 2005. "Investigating Intercollocations – towards an Archaeology of Text", *International Journal of Lexicography* 18(1): 1–24. doi:10.1093/ijl/eci001

Paltridge, B. 2006. *Discourse Analysis. An Introduction.* London, New York: Continuum.

Siepmann, D. 2005. "Collocation, Colligation and Encoding Dictionaries. Part I: Lexicological Aspects", *International Journal of Lexicography* 18(4): 409–443. doi:10.1093/ijl/eci042

Sinclair, J. 1991. *Corpus, Concordance, Collocation.* Oxford: Oxford University Press.

Sinclair, J. 2000. "Lexical Grammar", *Darbai ir dienos* 24: 191–203.

Sinclair, J., *et al.* 2004. *English Collocation Studies: The OSTI Report.* London, New York: Continuum.

Stubbs, M. 2001. "Texts, Corpora and Problems of Interpretation: A Response to Widdowson", *Applied Linguistics* 22(2): 149–172. doi:10.1093/applin/22.2.149

Tognini-Bonelli, E. 2001. *Corpus Linguistics at Work.* Amsterdam/Philadelphia: John Benjamins Publishing Company.

Vystrčilova, R. 2000. "Legal English", *Philosophica* 73: 91–96.

Wei, Y. 1999. *Teaching Collocations for Productive Vocabulary Development* [Accessed on 12 May 2009]. Available from Internet: <http://www.eric.ed.gov>.

Weisser, M. 2007. *N-grams, Collocations and Idiomatic Constructions* [Accessed on 02 May 2009]. Available from Internet: <http://ell.phil.tu-chemnitz.de/analysis/collocations.html>.

Widdowson, H. G. 2000. "On the Limitations of Linguistics Applied", *Applied Linguistics* 21/1: 3–25.

Sources

Judgments on appeal of the ECJ: <http://eur-lex.europa.eu/>.

BNC: <http://corpus.byu.edu/bnc/>.

Computational tools:

Scott, M. 2009. *WordSmith Tools Version 5.0*. Liverpool: Lexical Analysis Software.

Log-likelihood calculator. Available from Internet: <http://ucrel.lancs.ac.uk/llwizard.html>.

TEISĖS TERMINŲ KOLOKACIJOS EUROPOS TEISINGUMO TEISMO APELIACINIUOSE SPRENDIMUOSE

Daiva Macko

Darbo tikslas – nustatyti dažniausiai apeliaciniuose Europos Teisingumo Teismo sprendimuose vartojamas teisės terminų kolokacijas ir atrinktųjų kolokacijų vartoseną, palyginti su bendrinės anglų kalbos ypatybėmis. Straipsnyje vadovautasi metodologinėmis tekstynų lingvistikos nuostatomis. Buvo tirti veiksmažodiniai termino *Court* kolokatai. Bendriniu anglų kalbos šaltiniu pasirinktas Britų nacionalinis tekstynas (BNC). Kiekybinėje darbo dalyje pristatomi statistinės analizės rezultatai – įvertintas atrinktųjų kolokacijų statistinis reikšmingumas, gautų statistinių duomenų pagrindu kolokacijų vartosena teisinėje kalboje lyginama su bendrine anglų kalba. Kokybinėje tyrimo dalyje pateikiama tyrimo tekstyne dažniausiai vartotų pasirinkto sintaksinio modelio kolokacijų gramatinė ir semantinė klasifikacija. Tyrimo išvados patvirtina, kad statistinio reikšmingumo ir vartojimo dažnumo atžvilgiu kolokacijų vartosena Europos Teisingumo Teismo apeliacinių sprendimų anglų kalboje skiriasi nuo tų pačių kolokacijų vartosenos bendrinėje anglų kalboje. Tyrimo tekstyne nagrinėtos kolokacijos pasižymi ir savitomis semantinėmis ypatybėmis. Šie rezultatai parodo, kad specializuota teisinė anglų kalba, palyginti su bendrine vartosena, leksiniu požiūriu pasižymi reikšmingais skirtumais, į kuriuos mokant ir mokantis teisinės anglų kalbos yra labai svarbu atsižvelgti.

Reikšminiai žodžiai: koligacija, kolokacija, kolokatai, semantiniai modeliai, statistinis reikšmingumas.

Įteikta 2010-01-05; priimta 2011-03-25