



Vladislav Hrežo

University of Prešov

ORCID: 0000-0003-0090-3931

Exploring Phraseology in EU Legal Discourse

Eksploracja frazeologii w języku prawniczym Unii Europejskiej

Abstract

The aim of the present article is to showcase EU legal discourse as a unique phenomenon of supranational specialized communication and on the basis of authentic data analysis identify specific lexical items with a focus on multi-word expressions while considering their structure and function in the analysed text. The present analysis consists in researching a selected monolingual EU Directive in its English language version while using a mixed method approach. The results of analysis indicate that the EU Directive analysed in the presented structural and functional study contains a large proportion of multi-word expressions distinctive for legal language while adhering to the specific distributional patterns regarding the different structural and functional categories of lexical bundles. The article also gives an overview of contemporary scholars' research accomplished in institutional-legal discourse and translation.

Key words: *EU legal discourse, formulaicity, functional analysis, hybridity, legal language, legal texts, lexical bundles, multi-word expressions, phraseology, structural analysis.*

Abstrakt

Celem niniejszego artykułu jest ukazanie dyskursu prawniczego, jaki ma miejsce w Unii Europejskiej, jako niezwykłego zjawiska międzynarodowej komunikacji specjalistycznej i, na podstawie analizy rzeczywistych danych, wyszczególnienie specyficznych jednostek leksykalnych, kładąc przy tym nacisk na wyrażenia wieloczłonowe, w trakcie badań struktury i funkcji analizowanego tekstu. Ukazana analiza obejmuje badanie wybranej anglojęzycznej Dyrektywy Unii Europejskiej, przy użyciu metody podejścia mieszanego. Wyniki przeprowadzonych badań wskazują, że analizowana pod kątem struktury i funkcji Dyrektywa unijna zawiera znaczną ilość wyrażen wieloczłonowych charakterystycznych dla języka prawniczego, które nie zakłócają specyficznego układu syntaktycznego rozważanego z punktu widzenia występowania różnorodnych strukturalnych i funkcjonalnych kategorii leksykalnych. Ponadto, artykuł umożliwia wgląd we współczesne badania naukowe, jakie miały miejsce w sferze formalnego dyskursu prawniczego i translacji.

Słowa kluczowe: *dyskurs prawniczy Unii Europejskiej, formalność, analiza funkcjonalna, hybrydowość, język prawniczy, tekst prawniczy, jednostki leksykalne, wyrażenia wieloczłonowe, frazeologia, analiza strukturalna.*

Introduction

EU legal discourse represents a highly specialized discourse whose idiosyncratic linguistic-translational features have been explored by a number of researchers. (e.g. Gibová, 2010; Bednárová-Gibová, 2016, 2020; Biel, 2014; Klabal, 2019; Mori, 2018; Seracini, 2020; Sosoni, 2012; Trklja, 2017) EU legal discourse is distinctive for its specific terms and standardized form that have developed in the course of the last three decades. The specificity of EU legal discourse also consists in its multilingualism and hybridity. Texts produced by the institutions of European Union are expressed in twenty-four official languages. Such a multilingual translation routine has led to the formation of a brand-new and advanced variety of legal language, also referred to as 'Eurolect'.

From a discourse-analytical standpoint, institutional-legal discourse can be interpreted as 'hybrid, reproduced, mirror-image-like and horizontal' texts (Bednárová-Gibová, 2020) of which hybridity is one of the most distinctive features. (Bednárová-Gibová, 2017) With regard to the classification of EU discourse from the perspective of legal translation, Garzone as early as

2000 placed EU translation into the category of 'hybrid texts'. The author defines hybrid texts as "international instruments generated within the supranational multicultural environment, in particular EU legislation; [where] all language versions are authentic". (Biel, 2014a, p. 54) With regard to the concept of text hybridity, it was firstly proposed by Trosborg as far back as 1997 who is attributed with the introduction of a more sophisticated classification based on the compliance with cultural standards and who perceives EU institutional texts as hybrid political texts rather than legal ones. Biel with reference to the hybridity of EU discourse states that "EU law is a melting pot for national legal systems, languages and cultures" and further suggests that "EU language should be perceived as a multilingual legal language realised in distinct legal varieties of national languages with an interdependent conceptual system". (Biel 2014b, p. 337) Bednářová-Gibová aims to figure out the sources of that hybridity and its manifestations at various linguistic levels. (Bednářová-Gibová, 2017) Furthermore, the author also examines hybridity as one of the crucial factors in the process of creating transculture.

Crystal and Davy in reference with phraseology in legal language claims that this field has been commonly researched in regard to formulaicity, which has been observed as one of the most distinctive features of legal language. (Crystal, Davy, 1969) However, Biel reports that "legal phraseology has not been a popular topic in legal language studies" regardless of its high level of formulaicity which legal discourse is characteristic of. (Biel, 2018, p. 12) Nevertheless, the author points out to the changing tendency that has just recently started and that has been caused due to "the surge of interest in phraseology within corpus linguistics, which found its parallel in the legal domain...". (Biel, 2018, p. 12) Stanislaw Goźdz-Roszkowski and Gianluca Pontrandolfo in the introduction to their 2018 publication *Phraseology in legal and institutional settings* points to the fact that earliest studies of phraseology in legal language aimed attention at "those lexical items that displayed the highest degree of fixedness and repetition, i.e. binomials (e.g. *signed and delivered, act and omission*) and their extended versions: multinomials. The presence of this type of expressions in legal language was rightly perceived as one of the indicators of its formulaicity and standardisation, which can in turn lead us to understand the stylistic preferences in legal drafting". (Goźdz-Roszkowski, Pontrandolfo, 2018, p. 3). The authors emphasize the need to explore and analyse patterns of formulaicity, lexical bundles, in connection with legal discourse for the sake of achieving a deeper insight into legal language.

Another field of examination connected to the research in phraseology is terminology. In connection with terminology, Kjaer (Kjaer, 2007) points to “a strand of research which focuses on multi-word terms and collocations where at least one lexeme is a term”. Nevertheless, Pontrandolfo states that phraseology may operate as an element connecting the term and the text. (Pontrandolfo, 2015) “Phraseology should also be viewed through its close links with discourse. Seen from this perspective, phraseological research involves examining the organisation of language beyond the level of a sentence or a clause and focusing on larger linguistic units, such as conversational exchanges”. (Stubbs, 1983, p. 1). Goźdz-Roszkowski i and Pontrandolfo further point to a crucial function of legal phraseology in legal translation in the sense that “it is one of the discourse elements which mostly contribute to the naturalness of the translated text” and “which is one of the most difficult elements translators have to tackle in their job...”. (Goźdz-Roszkowski, Pontrandolfo, 2018, p. 4) Indeed, the authors claim that apart from terminology, legal translation is connected to an issue of conventions in phraseology. With regard to the course of near future research in phraseology, Goźdz-Roszkowski, and Pontrandolfo mention that “the link between legal phraseology and plain legal language” will pose a challenge “for both the academic and the professional communities”. Legal phrasemes are repeatedly regarded as elements which often complicate reading and understanding. (Goźdz-Roszkowski, Pontrandolfo, 2018, p. 4).

Lexical Bundles in Legal Language

Formulaic nature of general English phraseology with a focus on non-paremiological multi-word units has been eloquently discussed by Bednárová-Gibová and Jesenská. They provide us with a comprehensive definition of multi-word units/expressions which goes as follows: “a formulaic phrase or an utterance, consisting of several lexical constituents, ready-made in terms of its production, functioning as a syntactic and semantic whole whose meaning may be non-compositional”. (Bednárová-Gibová, Jesenská, 2019, p. 120). The term *multi-word unit* used to label phrasemes is reported to be preferred by “the natural language processing community, whereas [the term] *phraseological unit* seems to be the preferred term in the field of phraseology”. (Goźdz-Roszkowski, Pontrandolfo, 2018, p. 42) Apart from the term *multiword unit* and *phraseological unit*, Corpas Pastor presents several other terms used to when referring to a *formulaic sequence*, among them “multiword lexeme, multiword lexical unit, fixed expression, phrase figée, set

expression, set phrase”. (Corpas Pastor, 1996, p. 17) With respect to legalese, Biel calls attention to lexical bundles, which are “the most commonly researched multi-word units in the frequency-based corpus-driven approach”. (Biel 2018, p. 11) As we may see, references to lexical bundles may vary on the basis of a preference of different authors and are as follows: clusters, n-grams, chunks, lexical phrases, multi-word expressions/units/items/terms, multi-word lexemes, multi-word lexical units, fixed expressions, etc. Biber and Barbieri refer to lexical bundles as “building blocks in discourse” and “a kind of pragmatic ‘head’ for larger phrases and clauses, where they function as discourse frames for the expression of new information”. (Biber, Barbieri, 2007, p. 270) In addition, Biel states that lexical bundles provide common and recognizable frames “retrieved from memory which are filled in with new information...”. (Biel, 2018, p. 11) According to Biber and Barbier and Hyland, lexical bundles are merely recognized on the basis of frequency criterion. (Biber, Barbier, 2007; Hyland, 2008) In connection with the classification of lexical bundles, Biel provides us with the division based on two types of criteria – formal criteria, which focus on the length and structure of lexical bundles, and functional criteria:

“The length-based categorization takes into account a number of constituents in a bundle: if it contains three words, it is referred to as a 3-gram; if four words, a 4-gram; if eight words, an 8-gram. The structural categorization is based on the grammatical structure of lexical bundles, depending on whether they contain noun, verb or prepositional phrases and clause fragments” (Biel, 2018, p. 11).

Similarly, from the structural viewpoint, Trklja introduces lexical bundles as sequences usually consisting of three to six words. (Trklja, 2017) “These sequences are incomplete structural units both in semantic and grammatical terms and can be classified into different classes according to their functions” (Trklja, 2017). In addition, the author remarks that one of the most extensive research into the lexical bundles with regard to their textual function was accomplished by Nesi and Basturkmen. (Nesi, Basturkmen, 2006)

Roszkowski and Pontrandolfo proposes the categorization of tendencies in corpus research towards legal phraseology “(1) research into collocations; (2) research into routine formulae, (3) terminographically oriented studies, (4) cross-linguistic studies of phraseology, including translation, and (5) semantics of legal patterns”. (Goźdz-Roszkowski, Pontrandolfo, 2018, p. 12) However, with regard to lexical bundles Biel (2018) specifies that they “do

not fit the existing categorizations of legal phraseology” and therefore research into lexical bundles is classified under the second category – research into routine formulae. (Biel, 2018). “A traditional classification groups legal phrasemes, e.g., into: (1) multi-word terms, (2) collocations with a term and (3) formulaic expressions and standard phrases”. (Biel, 2018, p. 12) The group of ‘formulaic expressions and standard phrases’ is dealt with in Kjaer. (Kjaer, 2007) Biel in her 2014 publication *Lost in the Eurofog: Textual Fit of Translated Law* introduces a different classification that is arranged “specifically for the genre of legislation” and which “ranges from the global textual level to the local microlevel: text-organizing, grammatical and term-forming patterns as well as term-embedding and lexical collocations”. (Biel, 2014, p. 12) Nevertheless, the author calls attention to the fact that none of the above stated categorizations “embraces lexical bundles, which typically cut across all these categories, both structurally and functionally” and further adds that lexical bundles ought to be viewed as a category of its own, “as a distinct class of legal patterns in its own right, identified on the basis of frequency-based criteria (and thus incompatible with classifications based on other criteria)”. (Biel, 2014, p. 12) Nevertheless, the structures and functions of lexical bundles are presented and analysed in a separate section of the present article based on the analysis of a specific EU Directive No 1161/2019-EN version.

The Level of Formulaicity of EU Discourse

As early as 1983, Pawley and Syder conducted one of the first linguistic research of formulaic language. They claim that cognitive lexicon of language users “consists of holistically stored linguistic sequences” and “refer to these sequences as ‘lexicalized or institutionalized sentence stems’”. (Trklja, 2017) In addition, Trklja (2017) refers to Biber and Conrad who as far back as 1999 arranged a corpus-driven method as an approach for investigation of formulaic language. (Trklja, 2017) “The method, which was further elaborated in subsequent studies (e.g. Biber et al., 2004; Biber, 2009) focuses on the distribution of frequently recurring fixed sequences of words called *lexical bundles*”.

With reference to lexical bundles in legal discourse Biel (2018) mentions certain publications which aim attention at “how lexical bundles vary across English language legal genres in three legal systems: the EU, England and Wales, and the US, respectively”. (Biel, 2018) The publications are as follows: Jablonkai (2010), Breeze (2013) and Goźdz-Roszkowski (2011). Ja-

blonkai in her 2010 article *English in the context of European integration: A corpus-driven analysis of lexical bundles in English EU documents* accomplished research into the use of English in EU discourse “by investigating the most frequent word combinations in English documents issued by EU institutions”. (Jablonkai, 2010, p. 1) The research was based on “a mixed-genre corpus for ESP purposes and analyses the corpus of EU genres as a whole against the British National Corpus (BNC) (Sampler, Academic, News, Fiction sections) rather than against a reference corpus of a comparable genre, i.e. a UK legal/administrative corpus”. (Jablonkai, 2010, p. 1) The main objective of the research was to scrutinize lexical bundles in English language version of EU texts from the perspective of their functions and structures. Jablonkai concludes that the presented study demonstrates the high level of formulaicity of written English EU discourse in contrast to the reference corpora by applying lexical bundles with a high degree of frequency, suggesting that a considerable amount of EU text is comprised of formulaic patterns. (Biel, 2018)

“The EU corpus contains twice as many bundle types and six times as many tokens as the Academic prose section of the BNC; these rates are even higher compared to the fiction, news and general sections of the BNC. As for structural properties of EU bundles, bundles with noun phrases and prepositional phrases dominate the list (80%), but there is also an untypically high number of verb phrase bundles against the reference corpora”. (Biel, 2018, p. 12)

The Categorizations of Lexical Bundles

Goźdź-Roszkowski in his 2011 large-scale study *Patterns of Linguistic Variation in American Legal English. A Corpus-Based Study* offers comprehensive list of the structural types of the bundles in legal texts: (1) noun phrase with *of*-phrase fragment, (2) noun phrase with other post-modifier fragment, (3) prepositional phrase expressions, (4) verb phrase with passive verb, (5) verb phrase with active verb, (6) anticipatory *it* + verb phrase (usually passive), (7) adverbial clause fragments, (8) (verb/adjective+) *to*-clause fragment, (9) (verb phrase +) *that*-clause fragments. (Goźdź-Roszkowski, 2011, pp. 113-114) It should be noted, however, that the following list of the structural types is based on the American Law Corpus whose legal discourse may differ to some extent from that of the European Union. The author emphasizes that the above stated categories often occur in most genres. In

his study, the author proposes a proportional analysis of lexical bundles occurring throughout different structural categories that can be found in each of the selected genres. “For instance, the so-called “quantifier expressions”, i.e. bundles beginning with *all* (*all or any of, all or any part, all or any portion*, have been found only in contracts where they account for 6% of all the bundles”. (Goźdz-Roszkowski, 2011, p. 114) The author further suggests that the major character of lexical bundles in legal texts is phrasal and that more than 70% of all the bundles comprise of noun phrase expressions and prepositional phrases while bundles consisting of a verb phrase and dependent clause are likely to occur rarely. Concerning ‘noun phrase expressions’ the author states that those with ‘of phrase fragments’ represents more than 15% of all the lexical bundles comprising “a range of meanings in legal genres and in particular it is widely used to refer to various legal and other entities” (ibid). Apart from noun phrases, the author mentions bundles consisting of ‘adverbial clause fragments’ which were identified in the following legal genres that are of primary importance: legislation, contracts and judgments.

Besides the structural classification of lexical bundles the author proposes the other categorization, which is built upon the previous classifications as presented by Biber et al. (2004), Biber (2006) and Hyland (2008). “A framework for functional analysis of lexical bundles in legal texts was designed in such a way as to reflect the specificity and somewhat unique nature of legal data”. (Goźdz-Roszkowski, 2011, p. 117) From such a perspective, Goźdz-Roszkowski provides us with the following functional classification of lexical bundles: The first category encompasses bundles of legal reference and its seven subclasses: temporal bundles, location bundles, attributive bundles, participative bundles, institutional bundles, terminological bundles, and the last subclass of legal reference bundles, procedure-related bundles. The second major category presented by the author is text-oriented bundles which consists of these subcategories: causative/resultative bundles, condition bundles, clarification/topic elaboration bundles, focus bundles, framing signals, structuring bundles, transition bundles. The third major category is represented by stance bundles which are further divided into epistemic and attitudinal (ibid).

Methodology of the Present Analysis

Despite recent trends in institutional translation focusing on sociological lines of enquiry (e.g. Koskinen, 2017; Dam, Zethsen 2014; Bednárová-Gibová, Madoš, 2019), this article draws essentially on a text-based appro-

ach. The approach to the analysis presented in this article was inspired by the methodology incorporated by Goźdź-Roszkowski (2011) in his corpus-based study *Patterns of Linguistic Variation in American Legal English*. The author argues that he adopts a comparable methodology that had already been applied in the earlier studies of lexical bundles by Biber (2006), Cortes (2004) and Hyland (2008). In spite of the fact that the structural and functional categorizations provided by Goźdź-Roszkowski were arranged for the purpose of the corpus-based study of American Legal English, I assume that it can be applied to European Legal English as well since the present analysis deals with the same legal genre – legislation. Nevertheless, I suppose that both American and European Legal languages share certain features and have some points of intersection in terms of using the identical lexical items in that specific and specialized field of language.

However, unlike the above-mentioned comprehensive study by Goźdź-Roszkowski, the present analysis serves only as a pilot analysis aiming attention at only one genre of legal language – legislation. For the purpose of the analysis, a random act of EU legislation – Directive (EU) 2019/1161 selected from the EUR-Lex database. The focus of the analysis is on multi-word expressions that consist of three and more words which are supposed to be indicative or characteristic of legal language occurring in a genre of legislation that comprises 'directives', 'regulations', 'decision', 'recommendations' and other legislative acts. The multi-word expressions that have been chosen for the purpose of the present analysis will be tested in further research with the help of corpus linguistic software tools and online resources providing phraseological information. Among such online resources are the following ones: IATE, JURITERM, Euroterm, TERMIUM Plus, and several others.

The objective of the article is to illustrate EU legal discourse as a one-of-a-kind instrument of supranational specialized communication and from the structural and functional viewpoint determine its specific lexical units. The present analysis consists in researching the monolingual EU Directive in its English language version while combining both qualitative and quantitative methods. Qualitativeness of the presented analysis consists in hypothesizing about the data obtained from the empirical approach towards the text analysis.

The present analysis addresses the following research questions:

- Which multi-word expressions that are characteristic of legal language can be identified in an EU legislative act?
- What is the structure of those multi-word expressions recognized in the chosen EU act?

- Which kind of multi-word expressions with respect to their structure and function occur in the analysed text most frequently?
- What are the functions of those expressions?

A Structural and Functional Analysis of Multi-word Expressions in an Act of Legislation

Distribution of multi-word expressions across different structural types

From a structural perspective, the total number of all the multi-word expressions identified in the act of legislation *Directive (EU) No 1161/2019-EN version* accounts for 439. A half of them were identified as prepositional phrases with the number of 225. The most frequent ones were of *the European Parliament and of the Council* occurring 23 times, *in order to* occurred 17 times, and *in accordance with* 16 times. Such a great proportion of prepositional phrases goes in line with the finding by Goźdź-Roszkowski claiming that “in legal texts, prepositional phrase expressions are the most common structure overall, comprising on average, one-third of all lexical bundles” (Goźdź-Roszkowski, 2011, p. 115) while admitting that some genres may contain even a larger portion of them. Prepositional phrases deliver a wide range of functions that are described in the second part of the analysis, however, the most frequent ones are time and place indication, and indication of a variety of “logical relations between prepositional elements”. (Goźdź-Roszkowski, 2011, p. 115) The second highest proportion of occurrence in the text of EU Directive is represented by noun phrases with *of*-phrase fragments with the total number of 93, which is almost a quarter of all the multi-word expressions. This finding is similar to that of Goźdź-Roszkowski (2011) who reports that the noun phrases with *of*-phrase fragments make up for more than one quarter of all lexical bundles identified in his study.

Concerning the functions of this type of expression, they serve many purposes in legal genre, among them making a reference to different legal entities, in our case EU institutions; contracting parties; subject matters of different types of legal texts; and last, but not least, quantity. The third most extensive structural group of multi-word expressions is represented by verb phrases with modal verbs, especially those consisting of *should* and *shall*. Verb phrases with *should* accounts for 24 and verb phrases with *shall* make up for 16.

For the purpose of providing answers to the first, second, and third research questions, the most representative categories and their examples

identified in *Directive (EU) No 1161/2019-EN* version are presented underneath in the table:

Table 1: Overview of different structural types of multi-word expressions identified in the analysis

Categories of different structural types of multi-word expressions	Prepositional phrases	Noun phrases with <i>of</i> -phrase elements	Verb phrases with <i>should</i>	Verb phrases with <i>shall</i>	Other verb phrases with modals	Adverbial clause fragments	Verb + <i>to</i> -clause fragments	Verb phrase + <i>that</i> -clause fragments	Other
Number of items found	225	93	24	16	17	19	8	4	33

– prepositional phrase expressions:

in order to, in accordance with, on the basis of, in the context of, in this regard, at that point in time, in support of, in addition to, in the choice of, by means of, by an array of, with the requirements of, as a result of, in parallel with, within the time limit for, for the purpose of, during that period, by the aggregate of, by reason of, in so far as, in the case of, with respect to, at an earlier stage of, at the same time, by the following, by the Commission, by the contracts, under an obligation, due to the, in line with, under this directive, within the meaning of, by the scope of, for the provision of, in the next decade, in the coming years, for the period after, in this Directive, in compliance with, in favour of, in cooperation with, etc.

– noun phrases with *of*-phrase fragment:

the promotion of, the opinion of the, the procurement of, the achievement of, the objective of, the implementation of, the aggregate of, the purpose of, the categories of, the adjustment of, the renewal of, an array of, the allocation of, the inclusion of, the benefits of, etc.

– verb phrases with *should*

should be identified, should not be affected, should be aligned, should be defined, should not be blended, should also be counted, should not be included, should be noted, should be set, should be considered, should be provided, sho-

uld be required, should be facilitated, should be conferred on, should be accompanied, should therefore be amended, should take place, should take place, should set out, should be able to, should not apply to, should contribute to.

– verb phrases with shall

shall be defined, shall not be blended, shall be taken into account, shall be calculated, shall be assisted, shall be terminated, shall be presented, shall be laid down, shall apply to, shall not apply to, shall ensure that, shall continue to, shall submit to, shall bring into force, shall communicate to, shall enter into force.

– other phrases with modals

may adopt measures, can contribute to, can help ensure, can effectively help, should pursue policy measures, must be accelerated, will help to, will continue to, will contribute to, etc.

– adverbial clause fragments

as defined in, as laid down in, as compared to, as stated in, as noted in, as regards its scope, as described in, as reported in, as set out in, as was announced in, etc.

– verb + *to*-clause fragment

to exempt from the requirements, to meet the commitments, to meet the targets, to increase the proportion of, to amend that Directive, to implement this Directive, to comply with this Directive, to ensure uniform conditions.

– verb phrase + *that*-clause fragments

the Commission announced that, Market forecasts estimate that, the Commission concluded that, Member States shall ensure that.

– noun phrases connected to different entities, EU institutions, contracting parties, subject matters of legal texts, etc.

the Member States, the European Parliament, contracting authorities and contracting entities, the procurement procedures, the United Nations Economic Commission for Europe, etc.

Distribution of multi-word expressions across different functional types

The second part of the analysis is intended to answer the third and fourth research question. As can be inferred from the overview, multi-word expressions of the different functional categories are distributed in the analysed EU Directive unevenly. The most represented functional category of lexical bundles is that of 'legal reference' accounting for 239 items found while the least represented is the category of 'stance bundles' with only 36 items identified. Such a distribution across these functional categories corresponds

to that of Goźdz-Roszkowski claiming that “legislation, contracts and professional articles are clearly focused on legal reference, with relatively few bundles expressing stance and a moderate number of text-oriented bundles”. (Goźdz-Roszkowski, 2011, p. 118) Such an adherence of lexical bundles to that distributional pattern observable in legal texts is a very interesting phenomenon.

’Legal reference’ bundles refer directly to both abstract and physical legal issues concerning entities, contracting parties, institutions, instruments, concepts, subject matters, processes, etc. ’Text-oriented bundles’ are applied in order to signal both relationships between different textual segments and discourse organization. ’Stance bundles’ serve for expressing attitudes or assessments. (Trklja, 2017; Goźdz-Roszkowski, 2011)

Table 2: Overview of multi-word expressions across major functional categories

Functional categories	Legal reference bundles	Text-oriented bundles	Stance bundles
Number of items found	239	90	36

Legal reference bundles

Table 3: Overview of subcategories of legal reference bundles based on the number of their occurrence

Categories of legal reference bundles	Temporal reference bundles	Location bundles	Attributive bundles	Participative bundles	Institutional bundles	Procedure-related bundles
Number of items found	16	42	96	7	70	8

Temporal bundles

The first subclass of the major category of legal reference bundles are time-indicating, which is typical for the genre of legislation to use time

expression frequently. From the structural viewpoint, temporal bundles are usually prepositional phrases. (Goźdz-Roszkowski, 2011) They can be exemplified as follows:

at that point in time, within the time limit for, at an earlier stage of, in the coming years, for the period from, for the period after, for the first reference period, etc.

Example: *The impact assessment recognised that markets for low- and zero-emission urban buses are characterised by increased market maturity, whereas markets for low- and zero-emission trucks are **at an earlier stage of market development.** (Directive (EU) No 1161/2019-EN version, (11))*

Location bundles

The purpose of location bundles is to refer to places, different entities, EU institutions, contracting parties, subject matters of contracts, etc. *in the Official Journal of the European Union, in the Union, in the contract, in Member States, in this Directive, etc.*

Attributive bundles

The purpose of attributive bundles is to “describe legal entities, concepts, instruments and processes by specifying their attributes”. (Goźdz-Roszkowski, 2011, p. 122-123)

the opinion of the, the procurement of, the achievement of, the objective of, the implementation of, the provision of, the date of, the purpose of, the adjustment of, the renewal of, etc.

Example: *Such measures include the increased use of Union funds to support **the renewal of public transport fleets and better exchange of knowledge and...** (Directive (EU) No 1161/2019-EN version, (26))*

Quantifying bundles

one or more, part of a, the proportion of, the number of, a percentage of, the total number of, the aggregate of

Example: ***The number of such vehicles, however, is expected to increase in the coming years, especially plug-in hybrids.** (Directive (EU) No 1161/2019-EN version, (13))*

Participative bundles

The role of participative bundles is to indicate “the presence and participation of various legal or natural persons in a particular legal process or

transaction. They help identify individuals or corporations that may benefit or, more generally, be affected by legal provisions or transactions or other actions". (Goźdz-Roszkowski, 2011, p. 124)

in support of, for the provision of, in favour of, etc.

Example: *'Directive 2009/33/EC of the European Parliament and of the Council of 23 April 2009 on the promotion of clean road transport vehicles in support of low-emission mobility'; (Directive (EU) No 1161/2019, Art. 1 (1))*

Institutional bundles

The function of institutional bundles is to make a reference to various legal entities, EU institutions, contracting parties, legal instruments, etc. In the genre of legislation there is a very frequent use of this type of bundles. (Goźdz-Roszkowski, 2011, p. 124)

the Member States, the European Parliament, the United Nations Economic Commission for Europe, the European Economic and Social Committee, etc.

Procedure-related bundles

The category of procedure-related multi-word expressions is the last from the extensive group of legal reference bundles. In connection with the use of 'shall', Goźdz-Roszkowski clearly illustrates several different functions of its using in the legal genre and the way a speech act is affected by such a usage of *shall*. "In many procedure-related bundles, a speech act is explicitly marked by the presence of shall. It is usually associated with carrying the illocutionary directive force of expressing command, imposing obligation or conferring legal rights". (Goźdz-Roszkowski, 2011, p. 127)

Procedure-related bundles with shall

shall apply to, shall not apply to, shall ensure that, shall continue to, shall submit to, shall bring into force, shall communicate to, shall enter into force

Example: *This Directive shall enter into force on the twentieth day following that of its publication in the Official Journal of the European Union. (Directive (EU) No 1161/2019-EN version, Art. (3))*

Text-oriented bundles

Table 4: Overview of subcategories of text-oriented bundles based on the number of their occurrence

Categories of text-oriented bundles	Causative/Resultative bundles	Condition bundles	Clarification/topic elaboration bundles	Focus bundles	Framing signals	Structuring bundles	Transition bundles
Number of items found	30	3	1	9	18	19	9

Causative/Resultative bundles

The first subclass of the second extensive category of lexical bundles indicates “causative connection between two propositions”. (Goźdz-Roszkowski, 2011, p. 131)

as a result of, by reason of, in order to, for the purpose of, for the provision of, due to the

Example: *Therefore, vehicles that meet the clean vehicles or zero-emission vehicles requirements **as a result of** retrofitting should also be counted towards the achievement of the respective minimum procurement targets. (Directive (EU) No 1161/2019-EN version, (15))*

Condition bundles

in the case of

Example: ***In the case of** vehicles using liquid biofuels, synthetic and paraffinic fuels, those fuels shall not be blended with conventional fossil fuels; (Directive (EU) No 1161/2019-EN version, Art. 4 (4)-(b))*

Clarification/topic elaboration bundles

at the same time

Example: *...to put the Union on a path towards low-emission mobility and **at the same time** strengthen the competitiveness of the Union’s mobility eco-system. (Directive (EU) No 1161/2019-EN version, (4))*

Focus bundles

The role of focus bundles is to indicate the introduction of a new topic. *by the following*

Example: *Article 2 is replaced by the following. (Directive (EU) No 1161/2019-EN version, Art. 1 (3))*

Framing signals

Hyland (2008, p. 14) defines framing signals as bundles that “situate arguments by specifying limiting conditions”. Goźdź-Roszkowski states that framing signals “are also used to specify the conditions under which a statement can be accepted, working to elaborate, emphasize or compare aspects of an argument”. (Goźdź-Roszkowski, 2011, p. 135)

on the basis of, in the context of, within the context of, with respect to, in this regard, in the case of, outside the scope of, in so far as, within the meaning of

Example: *Regulation (EC) No 595/2009 of the European Parliament and of the Council of 18 June 2009 on type-approval of motor vehicles and engines with respect to emissions from heavy duty vehicles (Euro VI)... (Directive (EU) No 1161/2019-EN version, Art. 4 (5)-(***))*

Structuring bundles

Hyland (2008, p. 17) states that structuring bundles function in two different ways: “[they] help organize the text by providing a frame within which new arguments can be both anchored and projected, announcing discourse goals and referring to text stages” and “[they] point to other parts of the texts to make additional material salient and available to the reader in recovering the writer’s intentions”. (Hyland, 2008, p. 17)

as defined in, as laid down in, as compared to, as stated in, as noted in, as regards its scope, as described in, as reported in, as set out in, as was announced in

Example: *...a vehicle of category M3, N2 or N3 using alternative fuels as defined in points (1) and (2) of Article 2 of Directive 2014/94/EU of the European Parliament and of the Council... (Directive (EU) No 1161/2019-EN version, Art. 4 (4)-(b))*

Transition bundles

In accordance with Hyland (2008), Goźdź-Roszkowski (2011, p. 137) proposes that transition bundles “are used to establish additive or contrastive links between elements”.

in addition to, as well as

Example: ***In addition to** zero-emission vehicles, today there are few light-duty vehicles with air pollutant emissions of 80% or less of the current emission limits. (Directive (EU) No 1161/2019-EN version, Art. (13))*

Stance bundles

The role of stance bundles as assumed by Goźdź-Roszkowski is “to express different attitudes or assessments”. (Goźdź-Roszkowski, 2011, p. 117) Stance bundles are of two major classes – epistemic and attitudinal.

Table 5: Overview of subcategories of text-oriented bundles based on the number of their occurrence

Categories of stance expressions	Epistemic stance expressions marking the Source or Perspective	Attitudinal/Evaluative (impersonal)
The number of items found	33	3

Epistemic stance bundles

Epistemic stance bundles are generally employed in order to indicate an author’s commentary “on the knowledge status of the information contained in the following proposition. Such status can be expressed as certain, uncertain, probable, possible, etc”. (Goźdź-Roszkowski, 2011, p. 138) In the case of the present analysis, however, epistemic stance expressions are predominantly used to refer to a source of certain information or label the point of view from which information is provided.

in accordance with, with the requirements of, in line with, in parallel with, in compliance with, under an obligation, under this directive, under each contract, etc.

Example: *Those powers should be exercised **in accordance with** Regulation (EU) No 182/2011 of the European Parliament and of the Council. (Directive (EU) No 1161/2019-EN version, (30))*

Attitudinal stance bundles

The purpose of attitudinal stance bundles is to convey attitudes of an author “actions or events described in the following proposition”. (Goźdź-Roszkowski, 2011, p. 138)

it is appropriate to, it is crucial to, it is desirable for

Example: *Given the very limited market for double-decker buses and their specific design limitations, it is **appropriate** to apply, during the first reference period covered by this Directive, lower minimum procurement targets for zero-emission vehicles... (Directive (EU) No 1161/2019-EN version, (16))*

Limitation of the Present Analysis

First and foremost, the limitations of the presented pilot analysis consist particularly in not applying corpus linguistic tools designed for effective analysis of texts while providing identification and precise quantification of the occurrence of different lexical items. Secondly, because of not using the tools for the investigation of the legislative act, the analysis was limited to only one act as the thorough analysis performed manually was time demanding. However, Biel and Engberg in connection with a traditional routine in researching legal texts argue the following: “[the] research into legal translation has been predominantly qualitative, product-oriented and descriptive and very often involved a manual analysis of small samples”. (Biel, Engberg, 2013) Klalal points out the recent advancement in the way of approaching research in legal translation arguing that we may observe a change in the preference of both quantitative and mixed approaches while frequently incorporating corpus linguistics. Thirdly, figuring out the exact number of each and every relevant multi-word expression identified in the text based solely on the analysis of just one legislative act certainly does not provide us with entirely satisfactory and valid results.

Furthermore, scrutinizing one legislative act in order to figure out frequencies of occurrence of individual lexical items would not serve the purpose, either. Nevertheless, Klalal argues that it is not necessary to incorporate extensive corpora in order to study legal language. Similarly, Bhatia claims that “due to the formulaic and conservative nature of legal language even an analysis of a single act may render comprehensive and valid results and frequencies are usually of secondary importance”. (Bhatia-Vijay, 2004) Last, but not least, Klalal in connection with frequency and size of the corpus proposes a statement that should be also taken into account and which may be crucial in order to understand the specificity, intricacy, and delicacy of the genres of legal language:

“Thus, for certain types of phraseological units (e.g. term-embedding collocations or multi-word terms) frequency should not be the only criterion (cf. Biel 2014). In addition, if studying terminology and phraseology of a specific branch of law, a single

act may be the only source of law applicable to the branch in the respective jurisdiction, and creating a larger corpus would compromise the homogeneity of the data” (Klbal, 2019, p. 168)

Last, but not least, the other limitation of the present analysis may consist in investigating only one type of legal texts, which is in this case the selected EU Directive. For further research there is a need to incorporate several other types of EU legal documents, such as Regulations, Treaties, Agreements, Judgements, Opinions, etc. in order to exclude potential genre-driven specificities.

Discussion and Conclusion

In spite of the fact that the present analysis is not intended to serve as a comparative analysis, but rather a mixed method approach study, it would be useful to put in contrast both the present and Goźdz-Roszkowski’s analysis in order to provide the answer for the proposed research questions. From the comparison of the data obtained from the present analysis and those provided by Goźdz-Roszkowski, it can be concluded that the EU Directive analysed in the presented structural and functional study covers the great number of multi-word expressions which are characteristic of legal language while adhering to the specific distributional pattern concerning the different structural and functional categories of lexical bundles. This is a very interesting phenomenon that can be observed in legal texts. The other inference that can be drawn from the analysis is that one of the most distinctive features of legal language is its high degree of formulaicity and repetitiveness. The present analysis of the authentic legal text also confirms the statement by Biber in which he suggests that “lexical units that cut across grammatical structures” and “have identifiable discourse functions [...] are important for the production and comprehension of texts”. (Biber, 2006, p. 155)

From the conducted research it follows that the selected text genre still offers a number of both theoretical and practical issues to reflect on, among them the question whether there are some specific lexical bundles that are exclusively characteristic of EU institutional legal texts (apart from those referring directly to EU institutions, EU concepts, EU parties, etc.) It should be noted as well that over the last few years EU legal language has experienced considerable progress, also as a result of the development of specialised corpus linguistic tools that allow for extensive explorations of sizeable legal text corpora. In addition, when exploring a degree of investigation to which

the present state of research into lexical bundles has proceeded, it is apparent that there has been a great advancement in the course of the last decade. Such a remarkable progress is evident when comparing the recent studies (cf. Biel, 2018; Goźdz-Roszkowski, 2011, 2018) to the earlier ones, for instance Gibová (Gibová, 2010, p. 112-115) where lexical bundles were referred to by hypernym “lexical clichés”. Such a shift from the hypernymous designation to the sophisticated functional and structural system of lexical bundles indicates a huge step forward in lexical analyses of institutional texts, especially with regard to phraseological analyses.

Bibliography

- Bednárová-Gibová, K. (2016). Acquis communautaire as supranational legal texts and their interlingual reproduction across multilingual Europe. In: *Procedia – Social and Behavioral Sciences*, vol. 236, <http://www.science-direct.com/science/article/pii/S1877042816316913>, 01.09.2020.
- Bednárová-Gibová, K. (2017). “Translating from a Lingua Franca in the Setting of EU Translation”. In: *Rudn Journal of Language Studies, Semiotics and Semantics*. vol. 8, no. 1.
- Bednárová-Gibová, K. (2020). *Towards an Understanding of EU Translation*. 2nd revised edition. Prešov: Prešovská univerzita. Available at: <http://www.pulib.sk/web/kniznica/elpub/dokument/Gibova3>, 01.09.2020.
- Bednárová-Gibová, K., Jesenská, P. (2019). *Selected Chapters in English Lexicology. Part II: Phraseology and Word-formation*. Prešov: Vydavateľstvo Prešovskej univerzity.
- Bednárová-Gibová, K., Madoš, B. (2019). “Investigating Translators’ Work-related Happiness: Slovak Sworn and Institutional Translators as a Case in Point”. In: *Meta: Journal des traducteurs/Translators’ Journal*, vol. 64.
- Bhatia-Vijay, K., Langton, N. and Lung, J. (2004). “Legal Discourse: Opportunities and threats for corpus linguistics”. In: *Discourse in the Professions – A Corpus Linguistics Approach*, vol. 16.
- Biber, D. (2006). *University Language: A Corpus – Based Study of Spoken and Written Registers*. Amsterdam: John Benjamins Publishing.
- Biber, D. (2009). “A corpus-driven approach to formulaic language in English: Multi-word patterns in speech and writing”. In: *International Journal of Corpus Linguistics*, vol. 14.
- Biber, D. Barbieri, F. (2007). “Lexical bundles in university spoken and written registers”. In: *English for Specific Purpose*, vol. 26.

- Biber, D., Conrad, S. and Cortes, V. (2004). "If you look at...: Lexical bundles in university teaching and textbooks". In: *Applied Linguistics*, vol. 25.
- Biber, D., Conrad, S., (1999). "Lexical bundles in conversation and academic prose". In: *Language and Computers*, vol. 26.
- Biel, L. (2014a). "The textual fit of translated EU law: a corpus-based study of deontic modality." In: *The Translator*, vol. 20.
- Biel, L. (2014b). Phraseology in legal translation: A corpus-based analysis of textual mapping in EU Law. In: L. Cheng, K. K. Sin, and A. Wagner (eds.), *Ashgate Handbook of Legal Translation*. London: Ashgate Publishing.
- Biel, L. (2018). Lexical bundles in EU law: the impact of translation process on the patterning of legal language. In: S. Goźdz-Roszkowski, G. Pontrandolfo (eds.). *Phraseology in Legal and Institutional Settings: A Corpus-based Interdisciplinary Perspective*. Abingdon: Routledge.
- Biel, L. Engberg, J. (2013). "Research models and methods in legal translation". In: *Linguistica Antverpiensia*, vol. 12.
- Breeze, R. (2013). "Lexical bundles across four legal genres". In: *International Journal of Corpus Linguistics*, vol. 18.
- Corpas Pastor, G. (1996). *Manual De Fraseología Española*. Madrid: Editorial Gredos.
- Cortes, V. (2004). "Lexical bundles and student disciplinary writing: Examples from history and biology". In: *English for Specific Purposes*. vol. 23.
- Crystal, D., Davy, D. (1969). "The language of legal documents". In: *English for Specific Purposes*. vol. 23.
- Dam, H. V., Zethsen, K. K. (2014). Translators in international organizations. A special breed of high-status professionals? Danish EU translators as a case in point. In: C. Angelelli, (eds). *The Sociological Turn in Translation and Interpreting Studies*. Amsterdam/Philadelphia: John Benjamins.
- Garzone, G. (2000). *Legal Translation and Functionalist Approaches: A Contradiction in Terms?* Geneve: Actes.
- Gibová, K. (2010). *O preklade anglických právnych textov EÚ. Lingvisticko-translatologická analýza*. Prešov: Vydavateľstvo Prešovskej univerzity.
- Goźdz-Roszkowski, S. (2011). *Patterns of Linguistic Variation in American Legal English: A Corpus-based Study*. Frankfurt am Main: Peter Lang.
- Goźdz-Roszkowski, S., Pontrandolfo, G. (2018). *Phraseology in Legal and Institutional Settings: A Corpus-Based Interdisciplinary Perspective*. London/New York: Routledge.

- Hyland, K. (2008). "As can be seen: Lexical bundles and disciplinary variation". In: *English for Specific Purposes*, vol. 27.
- Jablonkai, R. (2010). "English in the context of European integration: A corpus-driven analysis of lexical bundles in English EU documents". In: *English for Specific Purposes*, vol. 29.
- Kjaer, A.L. (2007). Phrasemes in legal texts. In: H. Burger, D. Dobrovolskij, P. Kühn, and N. R. Norrick (eds.), *Phraseology: An International Handbook of Contemporary Research*. Berlin: de Gruyter.
- Klbal, O. (2019). "Corpora in Legal Translation: Overcoming Terminological and Phraseological Assymetries between Czech and English". In: *Clina, Interdisciplinary Journal of Translation, Interpreting and Intercultural Communication*, vol. 5.
- Koskinen, K. (2017). "Going Localised – Getting Recognised. The Interplay of the Institutional and the Experienced Status of Translators in the European Commission". In: *Hermes – Journal of Language and Communication in Business*, vol. 22.
- Mori, L. (2018). *Observing Eurolects. Corpus analysis of linguistic variation in EU law*. Amsterdam: John Benjamins.
- Nesi, H., Basturkmen, H. (2006). "Lexical bundles and discourse signalling in academic lectures". In: *International Journal of Corpus Linguistics*, vol. 11.
- Pawlwy, A., Syder, F. H. (1983). Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. In: J.C. Richards and R. W. Schmidt (eds.), *Language and Communication*. London: Longman.
- Pontrandolfo, G. (2015). Investigating judicial phraseology with COSPE. A contrastive corpusbased study. In: C. Fantinuoli and F. Zanettin (eds.), *New Directions in Corpus-based Translation Studies, Translation and Multilingual Natural Language Processing (TMNLP)*. Berlin: Language Science Press.
- Seracini, F. L. (2020). *The Translation of European Union Legislation. A Corpus-based Study of Norms and Modality*. Milano: LED.
- Sosoni, V. (2012). "A Hybrid Translation Theory for EU Texts". In: *Vertimo Studijos*, vol. 5.
- The European Parliament and the Council. (2019). Directive (EU) 2019/1161. In: *Official Journal of the European Union, L 188/116*. <https://eur-lex.europa.eu/eli/dir/2019/1161/oj>, 01.09.2020.
- Trklja, A. (2017). A corpus investigation of formulaicity and hybridity in legal language: a case of EU case law texts. In: Goźdz-Roszkowski, S. and Pontrandolfo, G. (eds.). *Phraseology in Legal and Institutional*

Settings A Corpus-Based Interdisciplinary Perspective. London/New York: Routledge.

Trosborg, A. (1997). "Translating hybrid political texts." In: *Text Typology and Translation*. Amsterdam: Benjamins, vol. 26.

Correspondence concerning this paper should be addressed to Vladislav Hrežo, M.A. in English – research fellow at the Institute of British and American Studies at the University of Prešov (Slovakia).

E-mail: v.hrezo@gmail.com