

# Lexical Bundles in Contract Law Texts: A Corpus-Based Exploration and Implications for Legal Education

Abdullah Alasmary<sup>1</sup>

<sup>1</sup> Department of English Language & Translation, College of Languages and Translation, King Saud University, Riyadh, Saudi Arabia

Correspondence: Abdullah Alasmary, King Saud University, Riyadh 12372-7101, Saudi Arabia. E-mail: aasmary@ksu.edu.sa

Received: January 9, 2019      Accepted: February 3, 2019      Online Published: February 24, 2019

doi:10.5539/ijel.v9n2p244      URL: <https://doi.org/10.5539/ijel.v9n2p244>

## Abstract

This paper reports on a study which explores lexical bundles in Contract Law, a key subdivision of the legal discourse. Based on a corpus of full-length texts, a total of 117 patterns are retrieved, refined and further subjected to structural as well as functional analyses. The results show that text authors make use of a wide range of lexical bundles, most of which are structurally phrasal and functionally research-oriented. Text-structuring sequences and participant-oriented bundles appear in the corpus, but are comparably far less employed. Also, the analysis of data established the domain-specific nature of patterns which revolve around the concept of *contract*. This paper concludes by discussing these findings and their implications for language learning, teaching and the ESP/EAP pedagogy.

**Keywords:** legal education, contract law, lexical bundles, corpus analysis

## 1. Introduction

Several studies maintain that academic speech and writing involve the use of a large number of recurrent multiword constructions which can be located, retrieved and analyzed for their structural forms and discourse functions (Biber & Barbieri, 2007; Biber, Conrad, & Cortes, 2004; Biber, Johansson, Leech, Conrad, & Finegan, 1999; Cortes, 2004; Hyland, 2008a, 2008b). These patterns are studied using a range of terms, the most common of which is that of lexical bundles (e.g., Breeze, 2013; Durrant, 2017; Esfandiari & Barbary, 2017). Lexical bundles are perceived as “words which follow each other more frequently than expected by chance, helping to shape text meanings and contributing to our sense of distinctiveness in a register” (Hyland, 2008b, p. 4). The pervasive use of such lexical bundles is not restricted to a particular genre, register or discipline, as evidence shows that domains of various types and dissimilar communicative purposes employ a wide range of structurally different and functionally distinct bundle types.

The prominent status of lexical bundles has led researchers to examine these patterns in the writings of novice and expert authors (Cortes, 2004; Hyland, 2008a), students with different first-language backgrounds (Ädel & Erman, 2012; Bychkovska & Lee, 2017; Chen & Baker, 2010), and native- and non-native-English speaking professionals (Esfandiari & Barbary, 2017; Pan, Reppen, & Biber, 2016; Pérez-Llantada, 2014). Other studies took a more register-focused approach, investigating lexical bundles in contexts such as the European documents (Jablonkai, 2010), law (Breeze, 2013), mathematics (Cunningham, 2017), pharmacology (Grabowski, 2015) and psychology (Esfandiari & Barbary, 2017).

Despite the prevailing view that lexical bundles are the “basic building blocks of discourse” (Biber, Conrad, & Cortes, 2004, p. 371), not much research is done on the distribution of such bundles in the legal domain. This study addresses this gap, exploring the presence of lexical bundles in a collection of texts on Contract Law, a key subdivision of the legal discourse. There is solid evidence which points out that disciplines, even closely related ones, exhibit a greater variation in terms of the linguistic conventions commonly employed by the scientific communities engaging in the construction and dissemination of knowledge typical to these disciplines (Durrant, 2017). A purpose-customized corpus of written academic texts is thus built with the aim of uncovering the most frequently recurrent and widely dispersed patterns. A study as such is hoped to shed light on the structural as well as the functional features characteristic of the legal domain in general and in contract law in particular.

## 2. Academic Writing and Lexical Bundles Research

Written academic texts have been the subject of several studies aimed at unveiling their rhetorical structures, linguistic features and their communicative purposes. Biber et al. (2004, p. 374) argue that textbooks and classroom teaching are “arguably the two most important registers in the academic lives of university students”. Hyland (2009) maintains that “textbooks are indispensable to academic life, facilitating the professional’s role as a teacher and constituting one of the primary means by which the concepts and analytical methods of a discipline are acquired” (p. 68). Surveying the differences as well as the similarities that exist across registers, genres and styles, Biber and Conrad (2009) concluded that written textbooks are produced to inform and educate rather than to disseminate fresh ideas. The communicative focus of textbooks is usually placed on laying out well-established facts, rather than announcing previously unknown findings.

Academic textbooks have been studied for the use of academic bundles in a range of different contexts. Grabowski (2015) conducted a study in which a corpus of information leaflets, product summaries, clinical trial protocols and chapters from textbooks is examined for both keywords and lexical bundles. The results indicate that textbooks have the greatest number of keywords but the fewest number of lexical bundles compared with the three other sub-corpora. The concentration of a large number of keywords in this corpus, the author argues, is related to the discipline-specific nature of textbooks. The paucity of lexical bundles in textbooks, however, is attributed to the nature of texts which are lexically dense and less formulaic. In a similar study, Breeze (2013) explored the distribution of lexical bundles across four legal sub-registers: academic law, case law, legislation and documents. Drawing on a corpus of two million words, the researcher carried out a structural and functional analysis of repeated formulaic patterns unveiled as a result of the corpus analysis. Legislation and documents corpora manifest the widest range of bundles whereas academic and case laws include the fewest bundles. Structurally, the author adopts a lexico-grammatical approach, thus dividing bundles into four categories: content noun phrases, prepositional phrases, adjectival phrases and fragments containing a verb phrase. With the exception of case law, the greatest number of bundles in the three other register types involves content noun phrases denoting agents, institutions, and documents. Most bundles in academic law refer to either abstract or action entities. In a similar fashion, the corpus incorporating academic textbooks has the smallest range of bundles in a study conducted by Biber and Barbieri (2007) who contrasted the presence of such bundles across a wide range of registers and academic domains. Functionally, textbook corpus is dominated by, first, referentials, and then discourse-organizers. Stance expressions are the least employed bundle type.

The use of lexical bundles by nonnatives/novices has been contrasted against the use of the same bundles by native/expert writers with inconsistent, and to some extent contradictory, results (Ädel & Erman, 2012; Bychkovska & Lee, 2017; Chen & Baker, 2010; Cortes, 2004; Esfandiari & Barbary, 2017; Llanes & Muñoz, 2009; Pan et al., 2016). While some studies maintain that native and professional writers demonstrate a thorough understanding of a wider range of different recurrent patterns than do nonnatives and less experienced writers (e.g., Ädel & Erman, 2012), some other studies point to the opposite, that is, student- or novice-produced writings incorporate a great number of lexical patterns when compared with writings produced by natives or professionals (e.g., Bychkovska & Lee, 2017). These discrepancies arise as a result of differences in the study design, the discipline under study and the type of genre that is investigated.

## 3. Overview of the Legal Discourse

Legal language has been the subject of several research studies throughout the past decades. Much research into the legal discourse revolves around the syntax and semantics of the legal prose, with a particular attention given to the challenges facing novices and non-experts in understanding the legal content. Statements of legal nature are relatively long, densely nominal and distinctly complex as they comprise archaic and semi-archaic forms (e.g., hereinafter), rare expressions (e.g., annul) and opaque formulae (e.g., corporate veil). Legal texts, furthermore, incorporate a great number of familiar terms carrying unfamiliar meanings (e.g., distress & find), passivized constructions, odd prepositional phrases, performative markers and a wide range of law-specific Latin-origin concepts (Cao, 2007; Haigh, 2015; Trosborg, 1997). Another reason that makes the legal text difficult to decode lies in the fact that legal language is “system-bound” in which “terms denoting concepts derive their meanings from a particular legal system” (Northcott, 2012, p. 218). In this case, a widely used legal term in a specific judiciary system may not have an equivalent term in another system. Vass (2017) adds another layer of difficulty which concerns the increasing number of law students and professionals who come from a non-English-speaking background in which legal concepts, terms and rhetorical conventions are learned and delivered in the students’ native language. The inherently complex nature of the legal writings has given rise to what is now known as the *Plain English Movement* (Hartig & Lu, 2014) which calls for embracing a far clearer, less archaic, and more reader-friendly writing style accessible to a wider base of readership.

Over the past few years, there has been a significant amount of research on topics related to the legal discourse from an ESP perspective. While the study of Vass (2017) focuses on verb hedges in a one-million corpus of journal articles, supreme court agreements and supreme court disagreements, thus concluding that lexical verbs serving a hedging function are more pervasive in journal articles than in the other two genres, the research by Cheng and Cheng (2014) attempts to investigate epistemic modality in a corpus of civil cases in Hong Kong and Scotland, revealing no differences between the two legal systems with respect to the distribution of epistemic expressions serving to signal a degree of probability and possibility. In a survey of existing pedagogical resources relevant to legal education, Candlin, Bhatia, and Jensen (2002) conclude that the writing materials available for the students on how to approach legal prose are not fulfilling a clear pedagogical purpose, thus failing to meet the learner's writing needs, ignoring advances in linguistics theory and practice and are mainly delivered in an inaccessible manner.

In a corpus-based attempt to draw a line between disciplines, Durrant (2017) maintain that law is closely aligned with history, politics and English, as the distribution of patterns show that they share a great number of similar lexical bundles. Law, however, uses a rather distinctive set of recurrent patterns when compared with other disciplines such as physics, food sciences and chemistry.

#### 4. Methodology

In this section, I will outline the corpus upon which this study draws. A discussion of the bundle selection and refinement will follow suit, focusing primarily on the criteria which have been applied while extracting bundles from the corpus and the measures taken to refine the set of bundles resulting from the corpus analysis.

##### 4.1 Study Corpus

A study corpus is created to elicit lexical bundles meeting predetermined frequency and distribution parameters outlined in Bundles Selection Criteria and Refinement Section below. Texts making up the corpus are pooled from a variety of contract law subtopics, such as mistakes in contract law, theory of contract law, the modern law of contract and Chinese contract law (see Appendix A for a full list of books). Sections removed prior to corpus treatment include the publication information, copy rights violations warnings, acknowledgements, appendices, references, footnotes, endnotes, and tables of figures, cases, and statutes. Although there is no way to ascertain the language background of authors, the fact that the text is published by a key publisher attests to these authors' expertise and scholarship. Table 1 gives a comprehensive description of the corpus used in this study.

##### 4.2 Bundle Selection Criteria and Refinement

Although the criteria for selecting bundles from a corpus of naturally-occurring language differ from one study to the other, there seems to be a general consensus among researchers that a target lexical bundle should contain a specific number of words, recur beyond a particular frequency threshold and should also appear across a predetermined number of texts making up corpus under scrutiny. Given the exploratory nature of this study, the length of the bundle, its frequency of occurrence and its distribution across the corpus subparts will determine the process of locating and extracting bundles. Another step to distill data will follow, thus removing overlapped and subsumed bundles. The *Cluster Function* in the software program *Wordsmith Tools* (Scott, 2016) is used to synthesize four-word bundles from the corpus and the *Concordance Function* is also employed to retrieve concordance lines needed to determine the meanings as well as the functions of selected bundles.

As for the length of the bundle, it is common practice in previous research to focus on four-word bundles, as three-word bundles are unmanageably greater in number and are sometimes embedded in four-word bundles. Lexical bundles of greater length, such as five-, six- and seven-word bundles, do exist but the rarity by which they occur makes them of little interest to researchers (Cortes, 2013; Esfandiari & Barbary, 2017). With respect to the frequency of occurrence, bundles are selected if they occur 40 times per million words, a normalized score corresponding to a raw frequency of 133. This conservative threshold (see, Esfandiari & Barbary, 2017; Pan et al., 2016) is to ensure that only bundles which recur frequently are selected for the analysis. The total number of bundles meeting the frequency criteria amounts to 150, all of which were copied into an excel sheet for further distilling of the data. The third step involves removing bundles occurring in at least five texts (25% of texts in the corpus). The impetus behind using such a specific minimum range score is to avoid patterns that are idiosyncratically typical of a text or author and since this study draws on a limited set of full-length texts, it is methodologically appropriate to include for analysis the types of bundles with greater tendency to occur across a range of such texts. A total of six bundles occurring in less than 25% of the texts are removed, thus reducing the overall number of bundles to 144.

By looking at the list of bundles resulting from applying the sequence and range criteria, it becomes clear that

there is much overlapping between bundles. Chen and Baker (2010) identified two types of overlapping: complete overlapping and complete subsumption. The bundles *principles of international commercial* and *of international commercial contracts* are two parts of the extended bundle *principles of international commercial contracts*. The two bundles share the same frequency and dispersion profiles. Both bundles are combined in a single string with the word *contracts* enclosed into two brackets: *principles of international commercial +(contracts)*. Complete subsumption occurs when “two or more 4-word bundles overlap and the occurrences of one of the bundles subsume those of the other overlapping bundle” (Chen & Baker, 2010, p. 33). Examples include patterns such as *to the terms of* and *the terms of the* which are similar except in two lexical items.

Another procedure involves removing bundles which refer to specific judiciary entities such as the British House of Lords and the supreme court of a particular state (e.g., Supreme Court of Michigan). Three such bundles are eliminated because they are extremely context-dependent (Chen & Baker, 2010). Bundles removed due to overlapping and context-dependency amount to 27, thus minimizing the number of bundles to 117.

Table 1. Corpus description

Key statistics	Study corpus
Number of texts	20
Tokens (running words)	3,519,882
Tokens used for word list	3,331,677
Types (distinct words)	52,279
type/token ratio (TTR)	1.57
Standardized TTR	1.41
Mean word length (in characters)	4.94
Word length std. dev.	2.79

## 5. Results

A general overview of items on the final list (see Appendix B) reveals some interesting aspects of the legal vocabulary characteristic of the Contract Law. Lexical expressions co-occurring with the word *contract* are unsurprisingly dominating the list, thus reflecting the topic-specific nature of this register. The recurrent use of the sequences such as *for breach of contract*, *in breach of contract* and *the breach of contract* mirrors a serious concern among the legal community of a possible failure from one or both parties to maintain the binding nature of contractual agreements. Other patterns co-occurring with the term *contract* discuss what constitutes a contract as a legal document: *term(s) of the contract*, *terms in consumer contract*, *contents of the contract* and *matter of the contract*. Another interesting pattern emerging from the data concerns the use of lexical bundles which transcend register boundary, thus occurring in distinct contexts. Expressions such as *in the case of*, *on the other hand*, *in the context of*, *in respect of the*, *on the basis of* do not seem to be tied to a specific register. In the following two sections, the structural forms of bundles as well as their discourse functions will be discussed with examples taken from the corpus.

### 5.1 Structural Patterns of Lexical Bundles

One objective of the current study is to account for the grammatical structures of lexical patterns emerging from the corpus analysis. Drawing on the framework developed by Biber et al. (1999), lexical bundles can be broadly classified into noun-based, preposition-based and verb-based groups, each of which can be further classified into subgroups (see Table 2).

Noun-based bundles fall into two groups: a noun phrase followed by an embedded *of-phrase* fragment or a noun phrase which takes other post-modifier fragments. There are thirty-five bundles beginning with a noun or noun phrase followed by a post-modifying *of-phrase* fragment. A second noun-based subcategory of bundles involves the use of a noun phrase with either a post-nominal clause fragment (e.g., *the fact that the*, *the way in which*) or a prepositional phrase fragment (e.g., *party to the contract*, *remedies for breach*). Yet a third noun-based subcategory consists of a noun head premodified by nouns, adjectives or both (e.g., *the parole evidence rule*, *the unfair contract terms*). The bundle *the contract and the* is the final pattern which does not seem to belong to any of the subgroups outlined above and is thus considered a fragment. The second major category of lexical bundles in the collection of texts on Contract Law contains forty-three preposition-headed lexical bundles, nearly half of which take an *of-phrase* fragment as a post-modifier (e.g., *for breach of contract*, *at the time of*, *in the case of*). Yet a third major structural group consists of lexical bundles comprising a verb component. Three verb-based bundles begin with anticipatory *It* followed by copular verb and then either an adjective (e.g., *it is clear that*, *it is*

important to) or a modal auxiliary with a passive verb (e.g., it should be noted). A non-anticipatory *It* construction is also noted in the data in which the pronoun is a substitute of an unambiguous noun (it was held that). Passivized verb construction is represented by five different patterns: *it was held that* +(the), *the contract was made*, *was held to be*, *to be found in*. Within the verb category, moreover, some bundles have a *to*-clause pattern, occurring bundle-initially as in sequences such as (*intention*) +*to create legal relations*, *to terminate the contract*, *to rely on the* or bundle-finally such as *an invitation to treat*. Another subgroup of verb-based bundles is composed of a noun phrase followed by an extended verb phrase: *the court held that* +(the), *the contract has been* and *the contract had been*. There are three lexical bundles beginning with *that* and followed by a noun or a *there-be* construction: *that the contract was*, *that there is no* and *that there was no*. Finally, the list consists of two bundle fragments with a verb component that cannot be included within any of the preceding groups. The distribution of bundles according to distinct grammatical categories is depicted in Figure 1.

Table 2. Structural patterns of LBs

Structural pattern	Structural sub-group	No. of bundles	% of the entire corpus	Examples
Noun-based	NP+ of-phrase	35	30 %	<i>the court of appeal</i> , <i>the house of lords</i>
	Other NPs	21	18 %	<i>party to the contract</i> , <i>contract for the sale</i>
Preposition-based	PP+ of-phrase	26	22 %	<i>For breach of contract</i> , <i>at the time of</i> , <i>in the case of</i>
	Other PPs	17	14 %	<i>On the other hand</i> , <i>on the basis that</i>
Verb-based	Passive verb construction	4	4 %	<i>the contract was made</i> , <i>to be found in</i>
	Anticipatory It +be/modal+ adjective/passive verb	3	2.5 %	<i>it is clear that</i> <i>it is important to</i>
	To-clause constructions	3	2.5 %	<i>intention</i> ) + <i>to create legal relations</i> , <i>to rely on the</i>
	That-constructions	3	2.5 %	<i>That the contract was</i> , <i>that there was no</i>
	Subject+ VP	3	2.5 %	<i>the court held that</i> +(the), <i>the contract has been</i> ,
	Others	2	2 %	<i>as we have seen</i> ,
Total		117	100%	

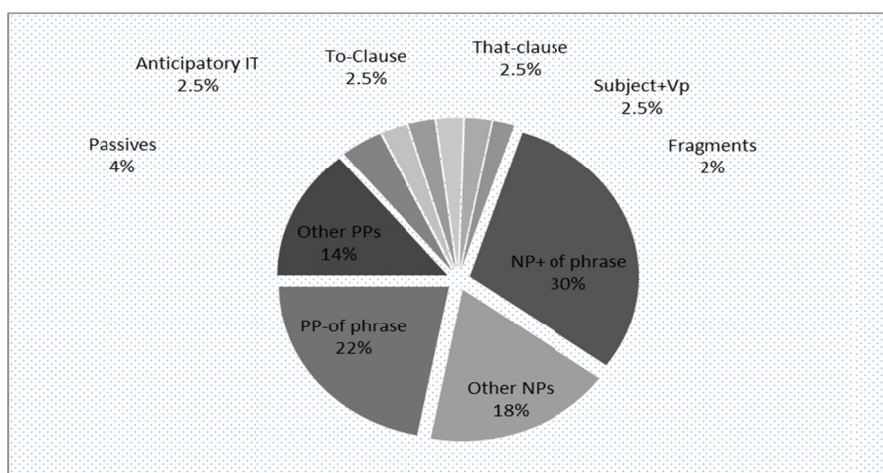


Figure 1. The distribution of structural patterns of lexical bundles

## 5.2 Functions

An important aim of the present study is to categorize bundles according to distinct functional groups. Hyland's (2008a, 2008b) framework helps to inform the classification process in this study where bundles fall into three major categories, namely *research-oriented*, *text-oriented* and *participant-oriented*. Although this functional framework is comprehensive enough to account for most bundles in the list, piloting with the data necessitates the following modifications:

- Hyland's *research-oriented*: *Location* sub-category is renamed as reference marker with three subheadings: time, place and agent markers. In this case, a bundle such as *the court of appeal* is used to refer to a specific entity whereas bundles such as *at the time of* and *the other party to* are used to refer to time and agent, respectfully.

- Following Biber et al. (2004), a new sub-heading titled *intangible framing attributes* is included within the research-oriented category, thus accounting for the functions served by bundles such as *the nature of the* and *the value of the*.

In some cases where functional boundaries blur, an inductive approach (Biber & Barbieri, 2007) is pursued, thus relying on the concordance lines in order to determine the function served by the target lexical bundle.

#### 5.2.1 Research-Oriented Bundles

As can be seen in Figure 2, bundles serving a participant-oriented function can be divided into sub-groups, each of which contains a number of distinct recurrent expressions. The greatest number of bundles are found in the topic-based category, whereas the smallest range of bundles occur in the description-based category.

##### 5.2.1.1 Time, Entity and Agent Markers

According to Hyland (2008b), research-oriented bundles “help writers to structure their activities and experiences of the real world” (p.13). Within this category, bundles can be used to mark time, place or entity. Reference markers alluding to time include two patterns: *at the time of* and *in the course of*. Bundles referring to a particular judiciary entity are represented by five bundles: *the court of appeal*, *the house of lords*, *by the house of*, *by the court of* and *of the court of*. The widest range of bundles in this sub-category are found to refer to agents. Examples include patterns such as *one of the parties*, *party to the contract*, *the other party to* and *the parties to the*. Here are examples from the data representing bundles serving to refer to time, entity and agent.

- “The contract was illegal *at the time of* its formation.” (time marker)
- “*The Court of Appeal* held that the creditor was bound to be consistent.” (entity reference marker)
- “It is possible for either both or only *one of the parties* to intend illegal performance.” (agent marker)

##### 5.2.1.2 Procedure

Several bundles in the list are found to help account for a specific procedure such as the ruling of a court or the intention of parties to enter into a contractual agreement. These include patterns such as *it was held that*, *the court held that* and *to create legal relations*.

- “*It was held that* it was unreasonable for the defendant to exclude liability for breach of both express and implied terms.”
- “The status of ‘*intent to create legal relations*’ has become disputed.”

##### 5.2.1.3 Description

The third research-oriented sub-category includes bundles used for describing a particular law-related action or legislation.

- “It would perhaps have been different had *the purpose of the* hire been specifically advertised in these terms.”
- “The mistake about *the application of the* Rent Acts was not a ground for declaring the lease void.”

##### 5.2.1.4 Intangible Framing Attributes

Some bundles within the research-oriented group tend to highlight the real or abstract nature of an entity. Bundles such as *the nature of the*, *the way in which* and *the value of the* help to exemplify the characteristics and qualities of a specific entity:

- “The lawyer must explain *the nature of the* transaction.”
- “The repairs would have cost twice *the value of the* ship.”

##### 5.2.1.5 Topic-Oriented Bundles

The largest group of bundles are domain-specific, that is, they are used to convey meanings typical of the contract law. Most of these domain-specific bundles revolve around the word contract; *for breach of contract*, *terms of the contract*, *the law of contracts*, *performance of the contract* and *matter of the contract*. A second set of domain-specific bundles serve to highlight some legislations such as *the statute of frauds*, *of the civil code*, *the uniform commercial code*, *the parole evidence law*.

- “*The terms of the contract* stated that the contract could be performed by the use of either of two named vessels.”
- “*The law of contract* is fundamental to any legal study.”



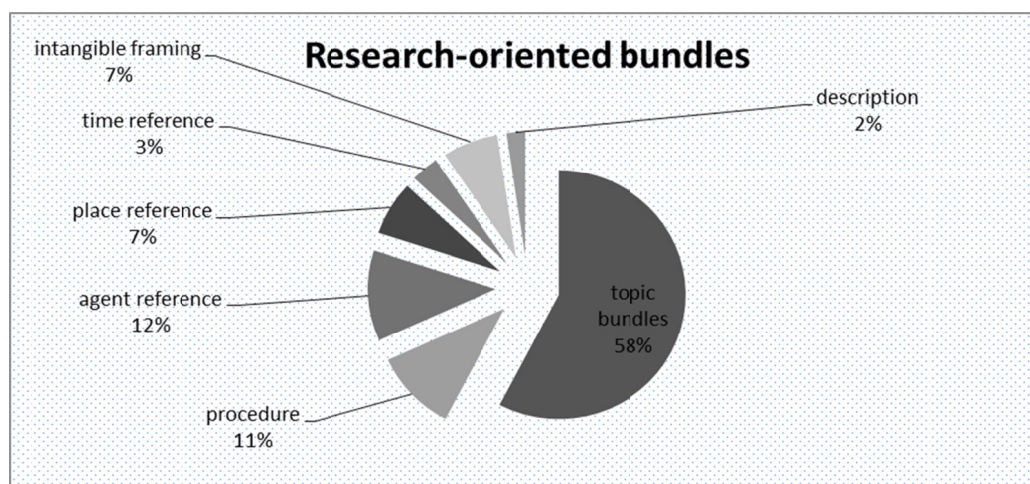


Figure 2. Distribution of research-oriented subgroups

### 5.2.2 Text-Oriented Bundles

As is shown in Figure 3, the second major functional category in the corpus-derived list contains bundles serving a text-organizing function. These expressions vary in the sub-function that they serve, but all seem to be “concerned with the organization of the text and its meaning as a message or argument” (Hyland, 2008b, p. 13). These bundles can be classified into the following sub-groups:

#### 5.2.2.1 Transition Signals

There are two bundles in the list which help authors to establish links between ideas presented in the text. The bundle *on the other hand* is used to draw contrast between two opposing ideas, opinions or arguments whereas the bundle *the difference between the* is employed to draw attention to the dissimilarities between elements.

- “An award of damages, *on the other hand*, brings the litigation to an end.”
- “This limited the loss to *the difference between the* overvaluation and the true valuation at the time of the contract.”

#### 5.2.2.2 Resultative Signals

A text-oriented sub-group is composed of three bundles serving as resultative signals, thus highlighting the causative relations existing between elements. The bundle *as the result of* showcases the outcome of an action being carried out whereas the two bundles *the effect of the* and *the benefit of the* are used to emphasize the consequences of a particular action or incidence.

- “As we have said, a party suffering loss *as the result of* a fraudulent misrepresentation can sue for damages in the tort of deceit.”
- “*The effect of the* listing was to cause the value of the property to drop.”

#### 5.2.2.3 Structuring Signals

A total of three bundles are used to “organize stretches of discourse or direct the reader elsewhere in text” (Hyland, 2008b, p. 14). The expressions *in relation to the*, *in accordance with the* and *to be found in* are found to act as structuring devices as is exemplified in the following statements:

- “It was *in relation to the* first issue that Benjamin's theory of mistake was argued by the plaintiffs.”
- “The Court of Appeal held that a contract had arisen, *in accordance with the* objective principle.”
- “Similar provisions are *to be found in* a number of international transport conventions.”

#### 5.2.2.4 Framing Signals

Some bundles in the list help “situate argument by specifying limiting conditions” (Hyland, 2008b, p. 14). The framing-signaling group of bundles consist of expressions such as *in the case of*, *on the basis of*, *in the context of* and *on the ground that*.

- “A term is transparent if it is expressed in plain and intelligible language and (*in the case of* a written term)

is legible.”

- “Restitution can apply in many legal contexts, but *in the context of* contract law, restitution will allow an injured party to recover money paid.”

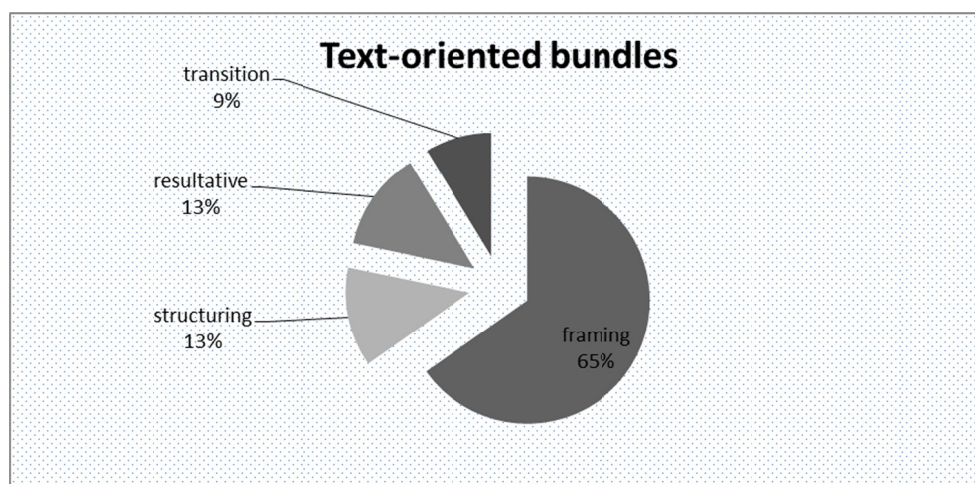


Figure 3. Distribution of text-oriented sub-groups

### 5.2.3 Participant-Oriented

The third major functional group is composed of bundles which serve either to mark the author’s stance or to engage the reader in the discussion that is being carried out. Stance expressions concern “the ways writers explicitly intrude into the discourse to convey epistemic and affective judgments, evaluations and degrees of commitment” whereas engagement bundles describe the “ways writers intervene to actively address readers as participants in the unfolding discourse” (Hyland, 2008b, p. 18). the distribution of bundles in the data is shown in Figure 4.

#### 5.2.3.1 Stance Expressions

A total of seven different bundles serve as stance markers, nearly three times greater than the number of engagement patterns. Some stance expressions revolve around the word *fact*: *the fact that the*, *the facts of the*, *on the facts of* while two others-*that there is no*, *that there was no*-employ negation to convey an epistemic meaning. Two other stance bundles, namely *it is important to* and *it is clear that*, are also used to showcase the writer’s opinion or judgment.

- “*The fact that the* transaction is executed rather than executory can be very relevant.”
- “So it contended *that there was no* real agreement.”
- “*It is important to* distinguish at the outset between two different types of mistake.”

#### 5.2.3.2. Engagement Bundles

Within the participant-oriented group, two bundles serve to engage the reader. An expression such as *as we have seen* is intended to draw the reader’s attention to a previous piece in the discourse whereas the bundle *it should be noted* +(*that*) is employed to alert the reader to the introduction of new information.

- “Traditionally, *as we have seen*, English law has refrained from imposing such a duty on parties to a contract or parties negotiating a contract.”
- “*It should be noted* that inclusion in the list does not mean that the term is unfair.”



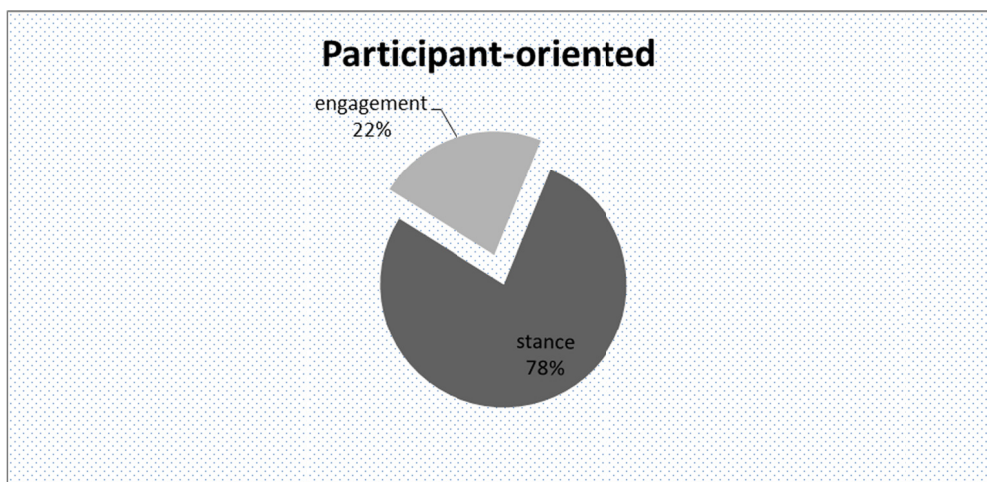


Figure 4. Distribution of participant-oriented sub-groups

## 6. Discussion

Despite decades of research into the distribution of multiword patterns in various disciplines, *Contract Law* seems to escape the attention of researchers. Drawing on methods from corpus linguistics (Biber et al., 2004, 1999; Hyland, 2008b) and genre analysis (Biber & Conrad, 2009), this study aims to bridge the gap in the existing knowledge on lexical bundle, thus seeking to extract, analyze and interpret these bundles in a corpus of full-length texts on *Contract Law*. By applying a set of predetermined frequency, dispersion and length criteria, a total of 117 bundles are obtained and their structural forms and functions are further analyzed. Of the whole list, twenty-nine bundles (nearly 30%) revolve around the concept of *contract*, an outcome that attests to the domain-specific nature of vocabulary terms (Biber & Barbieri, 2007). Such feature, furthermore, sheds doubt on attempts by some scholars (e.g., Liu, 2012; Simpson-Vlach & Ellis, 2010) to suggest discipline-transcending bundle lists that could be used by the students irrespective of their disciplines or study goals.

The structural analysis of lexical bundles reveals some interesting findings. First, nearly 90% of the expressions making up the list are phrasal whereas less than nine percent are verb-based. The greater use of noun-based patterns, on the one hand, and the paucity of verb-based ones, on the other, appear to confirm some previous results (Bhatia, 1993; Breeze, 2013) regarding the phrasal nature of the legal prose. Within the nominal group, the *NP/PP-of* structure is a particularly productive frame as it accounts for 50% of all bundles in the list. This grammatical device seems to be a feature of professional writing, as is maintained by Chen and Baker (2010) who found that this form is more abundantly used by expert authors than by native and nonnative student writers. The far less use of verb-based constructions seems to contradict some claims arguing that textbook writers rely more on recurrent verb patterns, given the fact that they are not constrained by time, like participants in a conversation, or space, like journal articles authors.

The analysis of bundles is extended to include the discourse functions that these bundles serve in the texts making up the corpus. It is unsurprising to find that a great number of patterns in the data serve a research-oriented function, an outcome that is consistent with similar ones cited in some previous research studies on academic discourse (Biber & Barbieri, 2007; Chen & Baker, 2010; Grabowski, 2015). Participant-oriented category is dominated by topic-based bundles. This finding goes against the one noted by Cortes (2004) who, using the term *referentials* instead of research-oriented, found out that few bundles in her analysis denote patterns specific to the disciplines of biology and history. Some discipline-specific patterns emerging from the analysis include reference markers which can be grouped further into three distinct categories: agent markers (e.g., party to the contract, the other party to), time markers (e.g., at the time of, in the course of), and place markers (e.g., by the house of, by the court of). Some of these patterns have been identified by Breeze, (2013) who used the term *content markers* as a labelling under which time, agent and place expressions are analyzed. The second largest functional group includes bundles serving as text-oriented markers, thus signaling to frame an opinion (e.g., in the context of), indicate transition (e.g., on the other hand), announce an outcome (e.g., as result of the) or to structure an argument (e.g., in relation to the). These text-structuring bundles are key characteristics of the legal text, as pointed out by Breeze (2013, p. 233) who maintains that "legal language is highly intertextual, and every new law, document or judgment establishes a web of connections to other laws,

documents and judgments”.

The third, and by far the smallest, functional group consists of participant-oriented bundles which can be further divided into stance and engagement markers, representing approximately 10% of all expressions unveiled in this study. The limited number of stance and engagement patterns in the list seems to give further credence to Bhatia's observation that legal language is “highly impersonal and decontextualized, in the sense that its illocutionary force holds independently of whoever is the ‘speaker’ (originator) or the ‘hearer’ (reader) of the document” (Bhatia, 1993, p. 188). The paucity of participant-oriented bundles can also be interpreted from a genre perspective, as written texts involve minimal interaction between the author and the reader (Biber & Conrad, 2009).

## 7. Implications

This study has key methodological and pedagogical implications. On a methodological level, future researchers will find the analytical frameworks adopted here easy to emulate while designing studies of similar goals. The steps for corpus compilation, extraction and refinement are thickly described in a way that allows for easier replication. It is also possible that items detailed here may be compared against similar ones elicited from texts of another disciplines (e.g., history, English) or texts of a similar sub-disciplines (e.g., common law, labor law). Studies as such are expected to deepen our understanding of the rhetorical practices shaping arguments in distinct as well as similar disciplines.

Pedagogically, this study has two important implications. Although the purpose of the current study is not to generate a definitive list of bundles in the contract law, it is hoped that language instructors, materials authors and textbooks compilers find some patterns in the list of greater benefit to their ESP/EAP students. In a short classroom activity, for example, students can be asked to examine the language of a legal contract with the help of the recurrent items in the list in order to determine how these items are functionally used to serve key communicative purposes. Another pedagogical implication is that instructors can draw on the corpus-derived examples outlined in the Findings Section while explaining the meanings as well as the functions of patterns in the list. In this case, learners not only have the opportunity to experience patterns as they occur in real contexts, but also can identify the different senses conveyed by each pattern based on real examples.

## 8. Conclusion

In conclusion, the role played by language in various academic setting is indisputably great, as is neatly encapsulated by Hyland who maintains that “educating students, demonstrating learning, disseminating ideas and constructing knowledge rely on language” (Hyland, 2009, p. 1).

The research reported here is an attempt to explore contract law, a key subdivision of the legal register, with the aim of unveiling recurrent multiword patterns. These patterns are then subjected to structural and functional analyses based on approaches and frameworks from corpus linguistics and genre studies. It is hoped that the findings as well as the discussion of these findings will increase our knowledge of the legal discourse in general and the law of contracts in particular.

## Acknowledgements

The author extends his appreciation to the Deanship of Scientific Research at King Saud University for funding this work through the research project No. R5-16-01-31.

## References

- Ädel, A., & Erman, B. (2012). Recurrent word combinations in academic writing by native and non-native speakers of English: A lexical bundles approach. *English for Specific Purposes*, 31(2), 81–92. <https://doi.org/10.1016/j.esp.2011.08.004>
- Bhatia, V. K. (1993). *Analysing genre: language use in professional settings*. London: Longman.
- Biber, D., & Barbieri, F. (2007). Lexical bundles in university spoken and written registers. *English for Specific Purposes*, 26(3), 263–286. <https://doi.org/10.1016/j.esp.2006.08.003>
- Biber, D., & Conrad, S. (2009). *Register, genre, and style*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511814358>
- Biber, D., Conrad, S., & Cortes, V. (2004). If you look at ...: Lexical bundles in university teaching and textbooks. *Applied Linguistics*, 25, 371–405. <https://doi.org/10.1093/applin/25.3.371>
- Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. Harlow, UK: Pearson.

- Breeze, R. (2013). Lexical bundles across four legal genres. *International Journal of Corpus Linguistics*, 18(2), 229–253. <https://doi.org/10.1075/ijcl.18.2.03bre>
- Bychkovska, T., & Lee, J. J. (2017). At the same time: Lexical bundles in L1 and L2 university student argumentative writing. *Journal of English for Academic Purposes*, 30, 38–52. <https://doi.org/https://doi.org/10.1016/j.jeap.2017.10.008>
- Candlin, C. N., Bhatia, V. K., & Jensen, C. H. (2002). Developing legal writing materials for English second language learners: problems and perspectives. *English for Specific Purposes*, 21(4), 299–320. [https://doi.org/10.1016/S0889-4906\(01\)00029-1](https://doi.org/10.1016/S0889-4906(01)00029-1)
- Cao, D. (2007). *Translating Law. Topics in Translation: 33*. Clevedon, UK: Multilingual Matters. <https://doi.org/10.21832/9781853599552>
- Chen, Y.-H., & Baker, P. (2010). Lexical Bundles in L1 and L2 Academic Writing. *Language Learning & Technology*, 14(2), 30–49.
- Cheng, W., & Cheng, L. (2014). Epistemic modality in court judgments: A corpus-driven comparison of civil cases in Hong Kong and Scotland. *English for Specific Purposes*, 33(1), 15–26. <https://doi.org/10.1016/j.esp.2013.07.006>
- Cortes, V. (2004). Lexical bundles in published and student disciplinary writing: Examples from history and biology. *English for Specific Purposes*, 23(4), 397–423. <https://doi.org/10.1016/j.esp.2003.12.001>
- Cortes, V. (2013). The purpose of this study is to: Connecting lexical bundles and moves in research article introductions. *Journal of English for Academic Purposes*, 12(1), 33–43. <https://doi.org/10.1016/j.jeap.2012.11.002>
- Cunningham, K. J. (2017). A phraseological exploration of recent mathematics research articles through key phrase frames. *Journal of English for Academic Purposes*, 25, 71–83. <https://doi.org/10.1016/j.jeap.2016.11.005>
- Durrant, P. (2017). Lexical Bundles and Disciplinary Variation in University Students' Writing: Mapping the Territories. *Applied Linguistics*, 38(2), 165–193. <https://doi.org/10.1093/applin/amv011>
- Esfandiari, R., & Barbary, F. (2017). A contrastive corpus-driven study of lexical bundles between English writers and Persian writers in psychology research articles. *Journal of English for Academic Purposes*, 29, 21–42. <https://doi.org/10.1016/j.jeap.2017.09.002>
- Grabowski, L. (2015). Keywords and lexical bundles within English pharmaceutical discourse: A corpus-driven description. *English for Specific Purposes*, 38, 23–33. <https://doi.org/10.1016/j.esp.2014.10.004>
- Haigh, R. (2015). *Legal English*. New York: Routledge. <https://doi.org/10.4324/9781315741840>
- Hartig, A. J., & Lu, X. (2014). Plain English and legal writing: Comparing expert and novice writers. *English for Specific Purposes*, 33, 87–96. <https://doi.org/10.1016/j.esp.2013.09.001>
- Hyland, K. (2008a). Academic clusters: text patterning in published and postgraduate writing. *International Journal of Applied Linguistics*, 18(1), 41–62. <https://doi.org/10.1111/j.1473-4192.2008.00178.x>
- Hyland, K. (2008b). As can be seen: Lexical bundles and disciplinary variation. *English for Specific Purposes*, 27(1), 4–21. <https://doi.org/10.1016/j.esp.2007.06.001>
- Hyland, K. (2009). *Academic discourse*. London: Continuum.
- Jablonkai, R. (2010). English in the context of European integration: A corpus-driven analysis of lexical bundles in English EU documents. *English for Specific Purposes*, 29(4), 253–267. <https://doi.org/10.1016/j.esp.2010.04.006>
- Liu, D. (2012). The most frequently-used multi-word constructions in academic written English: A multi-corpus study. *English for Specific Purposes*, 31(1), 25–35. <https://doi.org/10.1016/j.esp.2011.07.002>
- Llanes, À., & Muñoz, C. (2009). A short stay abroad: Does it make a difference? *System*, 37(3), 353–365. <https://doi.org/10.1016/j.system.2009.03.001>
- Northcott, J. (2012). Legal English. In P. Brian & S. Sue (Eds.), *The Handbook of English for Specific Purposes* (pp. 213–226). Chichester, UK: John Wiley & Sons. <https://doi.org/10.1002/9781118339855.ch11>
- Pan, F., Reppen, R., & Biber, D. (2016). Comparing patterns of L1 versus L2 English academic professionals: Lexical bundles in Telecommunications research journals. *Journal of English for Academic Purposes*, 21,

60–71. <https://doi.org/10.1016/j.jeap.2015.11.003>

Pérez-Llantada, C. (2014). Formulaic language in L1 and L2 expert academic writing: Convergent and divergent usage. *Journal of English for Academic Purposes*, 14, 84–94. <https://doi.org/10.1016/j.jeap.2014.01.002>

Scott, M. (2016). *WordSmith Tools* (7 ed.). Stroud: Lexical Analysis Software.

Simpson-Vlach, R., & Ellis, N. C. (2010). An academic formulas list: New methods in phraseology research. *Applied Linguistics*, 31(4), 487–512. <https://doi.org/10.1093/applin/amp058>

Trosborg, A. (1997). *Rhetorical strategies in legal language: Discourse analysis of statutes and contracts* (Vol. 424). Gunter Narr Verlag.

Vass, H. (2017). Lexical verb hedging in legal discourse: The case of law journal articles and Supreme Court majority and dissenting opinions. *English for Specific Purposes*, 48, 17–31. <https://doi.org/10.1016/j.esp.2017.07.001>

## Appendix A

### A Full List of the Textbooks That Make Up the Corpus Used in This Study

Andrews, N. (2015). *Contract law*. Cambridge: Cambridge University Press.

Ben-Shahar, O., & Porat, A. (2011). *Fault in American contract law*. New York: Cambridge University Press.

Cartwright, J. (2016). *Contract law: An introduction to the English law of contract for the civil lawyer*. Oxford: Hart Publishing.

Elliott, C., & Quinn, F. (2007). *Contract law*. Pearson Education.

Frey, M. A., & Frey, P. H. (2008). *An introduction to the law of contracts*. Clifton Park: Delmar Cengage Learning.

Furmston, M. P., & Cheshire, G. C. (2012). *Cheshire, Fifoot and Furmston's law of contract*. Oxford: Oxford University Press. <https://doi.org/10.1093/he/9780199568345.001.0001>

Heidemann, M. (2007). *Methodology of uniform contract law: The UNIDROIT principles in international legal doctrine and practice*. Berlin: Springer.

Hevia, M. (2012). *Reasonableness and Responsibility: A Theory of Contract Law* (Vol. 101). Springer Science & Business Media.

MacMillan, C. (2010). *Mistakes in contract law*. Bloomsbury Publishing.

McKendrick, E. (2005). *Contract law*. Basingstoke: Palgrave Macmillan.

Monaghan, C., & Monaghan, N. (2013). *Beginning contract law*. London: Routledge. <https://doi.org/10.4324/9780203569214>

Monahan, G. (2001). *Essential contract law*. Sydney: Cavendish. <https://doi.org/10.4324/9781843140979>

Neyers, J. W., Bronaugh, R., & Pitel, S. G. A. (2009). *Exploring contract law*. Oxford: Hart Pub.

O'Sullivan, J., & Hilliard, J. (2014). *Law of Contract*. Oxford: Oxford University Press. <https://doi.org/10.1093/he/9780199686933.001.0001>

Sefton-Green, R. (2009). *Mistake, fraud and duties to inform in European contract law*. Cambridge: Cambridge University Press.

Stone, R. (2009). *The modern law of contract*. New York: Routledge. <https://doi.org/10.4324/9780203876756>

Tepper, P. R. (2012). *The law of contracts and the Uniform commercial code*. New York: Delmar, Cengage Learning.

Turner, C. (2013). *Contract law*. New York: Routledge. <https://doi.org/10.4324/9781315879659>

Young, M. (2010). *Understanding contract law*. London: Routledge-Cavendish.

Zhang, M. (2006). *Chinese contract law: Theory and Practice*. Leiden: M. Nijhoff.

## Appendix B

### The List of Lexical Bundles Resulting from the Corpus Analysis

N	word	Normalized Freq.*	texts	Function
1.	the court of appeal +(held that the)/(that)/(in)	580	15	ROB**-place reference
2.	the house of lords +(in)	473	16	ROB-place reference
3.	for breach of contract	228	20	ROB-topic
4.	(to)+ the terms of the	226	20	ROB-topic
5.	at the time of +(the)	206	20	ROB-time reference
6.	in the case of	196	20	TOB***-framing
7.	on the basis of +(the)	180	20	TOB-framing
8.	the sale of goods +(act)	168	18	ROB-topic
9.	terms of the contract	162	19	ROB-topic
10.	the fact that the	154	20	POB****-stance
11.	on the other hand	146	20	TOB-transition
12.	for the sale of	139	20	ROB-topic
13.	(of)+the law of contract(s)	137	17	ROB-topic
14.	it was held that +(the)	136	15	ROB-procedure
15.	the court held that +(the)	123	20	ROB-procedure
16.	the nature of the	113	20	ROB-intangible framing
17.	(the) +performance of the contract	110	19	ROB-topic
18.	the subject matter of +(the)	110	19	ROB-topic
19.	of the contract law	109	6	ROB-topic
20.	the modern law of	105	9	ROB-topic
21.	on the basis that	103	16	TOB-framing
22.	in the context of	96	18	TOB-framing
23.	the statute of frauds	95	12	ROB-topic
24.	the value of the	93	19	ROB-intangible framing
25.	in relation to the	90	19	TOB-structuring
26.	of the contract and	89	20	ROB-topic
27.	to create legal relations	89	13	ROB-procedure
28.	by the house of	88	12	ROB-place reference
29.	by the court of	85	12	ROB-place reference
30.	as result of the	83	20	TOB-resultative
31.	(the)+ intention of the parties	83	20	ROB-topic
32.	on the ground that	83	16	TOB-framing
33.	on the part of	81	20	TOB-framing
34.	term of the contract	81	18	ROB-topic
35.	matter of the contract +(is)	80	17	ROB-topic
36.	one of the parties	79	19	ROB-agent
37.	in the course of	79	20	ROB-time reference
38.	that there was no	79	19	POB-stance
39.	of the subject matter	78	19	ROB-topic
40.	party to the contract	77	20	ROB-agent
41.	in the absence of	77	19	TOB-framing
42.	in the event of	75	18	TOB-framing
43.	of the civil code	73	6	ROB-topic
44.	remedies for breach of	73	19	ROB-topic
45.	mistake as to the	72	14	ROB-topic
46.	of the court of	72	14	ROB-place reference
47.	for the purpose of	69	18	TOB-framing
48.	of the contract the	69	19	ROB-topic
49.	that the contract was	69	18	ROB-topic
50.	contract for the sale	67	19	ROB-topic
51.	of the common law	66	17	ROB-topic
52.	the uniform commercial code	66	11	ROB-topic
53.	the intention of the	66	19	ROB-topic
54.	in the law of	65	15	TOB-framing
55.	the decision of the	65	16	ROB-description
56.	to the other party	65	20	ROB-agent
57.	the other party to	64	19	ROB-agent
58.	in this case the	64	20	TOB-framing
59.	the scope of the	64	18	ROB-intangible framing
60.	in breach of contract	63	18	ROB-topic
61.	the effect of the	63	17	TOB-resultative
62.	the contract was made	62	18	ROB-topic
63.	of the sale of	60	14	ROB-topic

64.	it is clear that	60	16	POB-stance
65.	(the)+part of the contract	60	18	ROB-topic
66.	the contents of the	60	14	ROB-topic
67.	the parole evidence rule	59	16	ROB-topic
68.	of the house of	59	12	ROB-place
69.	unfair contract terms act	57	12	ROB-topic
70.	in accordance with the	56	18	TOB-structuring
71.	of the other party	56	20	ROB-agent
72.	the parties to the +(contract)	56	19	ROB-agent
73.	it is important to	56	19	POB-stance
74.	damages for breach of	55	19	ROB-topic
75.	for the benefit of	55	18	TOB-resultative
76.	the purpose of the	55	19	ROB-procedure
77.	of the parties to	54	20	ROB-agent
78.	the existence of the	54	19	ROB-intangible framing
79.	the facts of the	54	19	POB-stance
80.	of the doctrine of	53	18	ROB-topic
81.	on the facts of	53	14	POB-stance
82.	contents of the contract	52	10	ROB-topic
83.	that there is no	50	20	POB-stance
84.	an invitation to treat	50	11	ROB-procedure
85.	the identity of the	50	17	ROB-intangible framing
86.	on the ground of	48	16	TOB-framing
87.	the way in which	48	17	ROB-intangible framing
88.	obligations under the contract	47	18	ROB-topic
89.	as we have seen	47	15	POB-engagement
90.	it should be noted +(that)	47	14	POB-engagement
91.	by the other party	46	17	ROB-agent
92.	in the light of	46	16	TOB-framing
93.	to terminate the contract	46	15	ROB-procedure
94.	the party in breach	45	13	ROB-agent
95.	the unfair contract terms	45	11	ROB-topic
96.	the doctrine of consideration	45	10	ROB-topic
97.	contracts for the sale	44	17	ROB-topic
98.	in respect of the	44	12	TOB-framing
99.	the breach of contract	44	17	ROB-topic
100.	the application of the	44	18	ROB-description
101.	the parties to contract	44	19	ROB-agent
102.	unfair terms in consumer	44	11	ROB-topic
103.	the doctrine of privity	43	11	ROB-topic
104.	was held to be	43	13	ROB-procedure
105.	rights of third parties	42	13	ROB-topic
106.	the use of the	42	20	ROB-procedure
107.	(of)+principles of international commercial	42	9	ROB-topic
108.	the contract has been	42	18	TOB-topic
109.	terms in consumer contracts	42	10	ROB-topic
110.	the difference between the	42	16	TOB-transition
111.	contracts rights of third	41	11	ROB-topic
112.	in the form of	41	19	TOB-framing
113.	the contract and the	41	19	ROB-topic
114.	to be found in	41	14	TOB-structuring signal
115.	the contract had been	40	18	TOB-topic
116.	to rely on the	40	17	ROB-procedure
117.	at the time the	40	17	ROB-time marker

Note. \*. Normalized Frequency (per million words)

\*\*. Research-oriented category

\*\*\*. Text-oriented category

\*\*\*\*. Participant-oriented bundles

## Copyrights

Copyright for this article is retained by the author, with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).