

# A Corpus-Based Approach to Studies in Legal Phraseology: An Overview

Ibrahim Bashir<sup>1</sup>, Kamariah Yunus<sup>1</sup> & Aliyu Abdullahi<sup>2</sup>

<sup>1</sup> Centre of English Language Studies, Faculty of Languages and Communication Universiti Sultan Zainal Abidin (UniSZA), 21300 Kuala Nerus, Terengganu, Malaysia

<sup>2</sup> Centre of Arabic Language Studies, Faculty of Languages and Communication Universiti Sultan Zainal Abidin (UniSZA), 21300 Kuala Nerus, Terengganu, Malaysia

Correspondence: Kamariah Yunus, Centre of English Language Studies, Faculty of Languages and Communication Universiti Sultan Zainal Abidin (UniSZA), 21300 Kuala Nerus, Terengganu, Malaysia. E-mail: kamariah@unisza.edu.my

Received: May 23, 2018      Accepted: June 20, 2018      Online Published: November 27, 2018

doi:10.5539/ijel.v8n7p50      URL: <https://doi.org/10.5539/ijel.v8n7p50>

## Abstract

The language of the law as it is called “legalese” has very distinctive lexical and structural patterns which in many ways different from the “traditional forms of language”. Its conservatism is linked directly to the need for unambiguous language that has already been tried and tested in the courts. By retaining to traditional lexis and structure lawyers can be confident that the language of the law is consistent and precise. This study aims to give some insights on apparent lexico-grammatical features characterised legal phraseology. The present study adopts a corpus-based approach to investigate those distinctive features of legal phraseology such as the uses binomial words, colligation of prepositions, prefabricated word combinations directly prescribed by law, and their semantic functions. This overview compiles data from the books, and empirical studies as well as theoretical and conceptual works conducted in the premises of legal phraseology. Some implications for English for specific purposes are given.

**Keywords:** binomial words, corpus-based approach, phraseology, legal phraseology

## 1. Introduction

Language is a system of meaning generation built up through the use of words and other linguistic forms like tones and emphasis. Thus, the notion of meaning is central to the study of the language and the most pervasive phenomena which intricate with diverse views of linguists. Many presuppositions in linguistics are essentially centred on meaning of words and range of references that are inherent in or related to or be part of meaning making such as context and discourse of use (Rashid, 2016; Rashid, Yunus, & Wahab, 2018). Notwithstanding, linguists have different views on language. The most pervasively difficult to explore are the aspects of *Semantics* and *Pragmatics* views on language. Halliday views language as a “system of meaning and that a grammar of a language is the study how meanings are built up through the use of words and other linguistic forms such as tones and emphasis” (Bloor & Bloor, 1995, p. 1). However, not all linguists view meanings in a direct way as Halliday did, there exist various perspectives that are divorced from the meanings. Some linguists for instance view language from its segmental structures (words, phrases and sentences) and how these segments work to form meanings. Words always play great roles in language studies of meaning variably as primary factors or secondary segments in more abstraction as a metaphoric or idiomatic expression. In the first half of the 20th century Chomsky from behaviourism perspective theorised language his popular theory (Universal Grammar). Chomsky (1981; 1986) fundamental concerns with linguistics competence in his three suppositions: what constitutes knowledge of language? How is such knowledge acquired? And hoe is such knowledge put to use? (White, 1989)

Later on inspiring revolutionist corpus linguistics john Sinclair (1991; 1994) postulates that only knowledge of word forms and grammar is not sufficient to express oneself idiomatically or naturally in a specific language. One requires to have the knowledge and linguistic skill to use language effectively and they it occurs naturally in a particular language discourse (Mitkov, 2017). This is the view of generative linguists which was in contrast with the early view of behaviourists and structuralists linguists. Pawley and Syder (1983) argue that, the

traditional approach the knowledge of language fail to account for nativelike fluency and study of language at it occurs naturally in a specific discourse. Modern linguists observe that language is indeed phraseological (Mitkov, 2017), that is to say language is communicated by thorough collections of lexical chunks which are naturally formulated and reused in a particular discourse. The meaning of such phrases can be presumed semantically and pragmatically in the context. Scholars, described these chunk of words using different terminologies, such as phraseological units (Naciscione, 2017), formulaic expressions (Ohlrogge, 2009), lexical bundles (Biber, Conrad, & Cortes, 2004), fixed expressions (Alexander, 1984; Rammell, Van Lancker, & Pisoni, 2018), set expressions (Shaikhullin & Alavi, 2015), multiword expressions (Mitkov, 2017), multiword units (Boers, Demecheleer, Deconinck, Stengers, & Eyckmans, 2017), idiomatic expressions (Bobrow & Bell, 1973; Cacciari & Glucksberg, 1991; Titone & Connine, 1999), colligations (Hoey, 2005; Yunus & Su'ad, 2016).

Moreover, phraseological expressions are central to a wide range of natural processing and applied linguistics, and they are interest of linguists in the last three decades, especially, the terminologists, translators, corpus linguists, lexicographers and researchers in language teaching and learning (Mitkov, 2017). They are important aspects of linguistics studies they constitute up to 45% of spoken English and up to 21% of academic prose in English (Biber et al., 1999). They are highly relevant in studies of English for specific proposes (ESP) and crucial elements in computational analysis of naturally occurring language (corpus linguistics), cognitive linguistics, pragmatics and discourse analysis studies. Their special roles in legal discourse are one of the motives of carrying this review. Legal language is characterised as a special discourse full of formulaic expressions, binomials (Yunus & Su'ad, 2016; Biel, 2016) whereas words play the roles of concepts.

The discussion set forth here is primarily concerned with an overview of a corpus-based approach to studies in legal phraseology. Notwithstanding, corpus linguists, computational linguists and other linguists who are interested in applying electronic corpora often have different views concerning the language, but corpus linguists generously paid a special attention “to not only with what ways, structures, or uses are possible in a language but also with what is probable—what is likely to occur in a language use” (Kennedy, 2014, p. 8).

## **2. An Overview of Studies on Phraseology**

### *2.1 Phraseology*

Howarth (1998) defines the term “Phraseologism” or “phraseology” as the co-occurrence of lexical items and one or more additional item(s) which together function as one semantic unit in a clause or sentence. According to Cowie (2002) Raisedmarie Glaser has contributed substantially to the present-day phraseological studies in her research work in which she investigates the roles of “multi-word units in special purpose language”; as well as problems in translating idioms and that of collocations. Her work reveals how heavily phraseology helps in understanding “stylistic expressiveness of the literary texts”.

Moreover, the system of phraseological categories is associated to Glaser, and developed in a number of theoretical and descriptive studies, draws on the analytical schemas initiated by the Russian phraseologists. However, it is more elaborate than most of other systems and especially notable for its treatment of a range of propositional (i.e. sentence-length) expressions, including proverbs, maxim, slogans, and quotations.

### *2.2 Phraseology in Specialised Languages (LSP)*

The term language for specialised purposes known in old tradition as specific purposes (LSP) had a long history that can be traced back to the time of Roman and Greek Empires (Dudley-Evans & St John, 2012), although it can be given different interpretations and applications. Bathia, Sanchez-Hernandez and Perez-Paredes (2011) posit that the current studies in LSP have increased and moved to a stage where it has twofold status as independent discipline with its own “research interest”, “methodology” and “applications” and as multi-disciplinary in the sense that it draws insight from other disciplines such as corpus-based analysis, critical discourse analysis, genre analysis and many other fields of linguistics. In addition, Gollin-Kies, Hall and Moore (2015) posit that the question of specificity in LSP is what poses a lot of challenges among the linguists. Thus, there is need to draw parameters to distinction between what has been called English for general academic purposes (EGAP) and English for Specific Academic Purposes (ESAP) (Hyland, 2006).

What do we mean when we talk about language for specific purposes? Is there something specific about the language to be used, and if there is in what way can that difference be defined? Does there exist, for example, a definable language specific to the teaching of physics, to the writing of business letters, to the writing of international contracts, to the verbal exchanges between diplomats? Does the difference, if it exists, resides in the specialised vocabulary used, in the different meanings ascribed to words in different contexts, in the different collocations, in the way that sentences and ideas are sequenced and combined to form a coherent communicative

event, in the interpersonal relations that are realised in the text, in the constraint of what can and cannot be said and on different roles of the event's participants, in the values that are embedded in the text? If there are differences, are they probabilistic or can rules be specified? Is it possible to prescribe or describe a one-to-one mapping between form and function?

According to Strecker, (1985) Language is basically pragmatic means of communication which we possess and it is a provisional and ever-changing stockpile of examples (Gollin-Kies et al., 2016). Secondly, some of these issues can be addressed by observing the phrase's constituents of the term; as the term "specific" implies the delimitation in what we intend to address. One of the striking issues in LSP is on how to determine the extent of specificity, for examples on basis of functionality, topic, discipline or genre, and a target situation. The question here is that how specifically do we need to be? Indeed about what? How much of what? Meanwhile, Johns (2006) claims that the specificity is the distinctive and the principle tenet of the Dudley and John (1998) specific courses need specific target situation and the individuals concerned and extensively used of authentic materials.

Moreover, modern phrenologists concerned with the study of language forms, structures and functions especially using the corpus-based method of analysis. The attention was shifted from Chomskian view of "decontextualisation" of sentences whereas analysis was carried out separately with no reference to any context to now consider the discourse that is language in its context of use (Flowerdew, 2013) and analysis of supra-sentential elements believing that language is more than "grammar and vocabulary". Perhaps, the emphasis of the modern linguistic studies was on the Lexico-grammatical elements of language and their functions at beyond sentential levels. Examples of languages for specific purposes: language of advisement, language of politics, legal language (legalese), language for religious purposes, businesses language and so forth.

### *2.3 Phraseology in English for Specific Purposes (PESP)*

English for specific purposes (ESP) is an umbrella term that encompasses varieties of English uses for various purposes, such as English for academic purposes (EAP), English for vocational, occupational and professional purposes (ECP/EOP/EPP) as well as many "finer categories" including English for business purposes (EBP) English for engineering purposes or technology related fields (EEP/ETP) and many other specialised areas where English is used to serve those specific purposes (Jordan, 1997; Paltridge & Starfield, 2013; Gollin-Kies et al., 2016). As a field of study ESP is a special course offered mostly to group of adults that need language to serve a particular communicative purposes involving different genres at different discourse, Paltridge and Starfield (2013) add that the users are often "homogenous in terms of learning goals" and in some cases heterogeneous in terms of language "proficiency".

The status of ESP as theory is debatable, Dudley-Evans (1997) argues that ESP was not established as an independent theory let say like "Communicative Language Teaching and Second Language Acquisition"; on the same vein Duldley-Evans and St John (2012) narrate that it is debated that "ESP lacks an underlying theory", however, in their opinion ESP has that theory that could be drew from either "the specific nature of the texts that learners requires knowledge of or on the basis of needs-related nature of the teaching", though most of the studies of ESP concentrate more on the procedures of ESP teaching and learning and also give more attentions to the course design to "learners" specific needs rather than on theoretical matters'. It could be believed that the definition of ESP should reflect certain features: it should be designed to saturate specific needs of the learner in a particular discourse, the methodology of teaching and learning should be different from that of general English, it should be centred on the use of language skills, grammar, lexis, register appropriate to the particular discourse or genre.

### *2.4 Classification of English for Specific Purposes (ESP)*

English for specific purposes (ESP) is one the most influential phenomena in ELT and applied Linguistics especially in the leading word power economics countries like USA, UK, Australia for the increase innovation in educational research and practice, business and rapid development in computer technology since 1950s and 1960s (Dudley-Evans & St John, 2012) and more interestingly the need of learning English for international communicative purposes across the different discourse and genres. ESP was basically classified into two main areas: English for Academic Purposes (EAP) and English for Occupational Purposes (EOP), these areas have sub-divisions which are mostly outlined in a diagram. Presenting this classification in the diagram is important (Dudley-Evans & St John, 2012) as it helps in mapping the degree and density of specificity which was useful in designing syllabus of ESP and carrying needs analysis in ESP discourses.

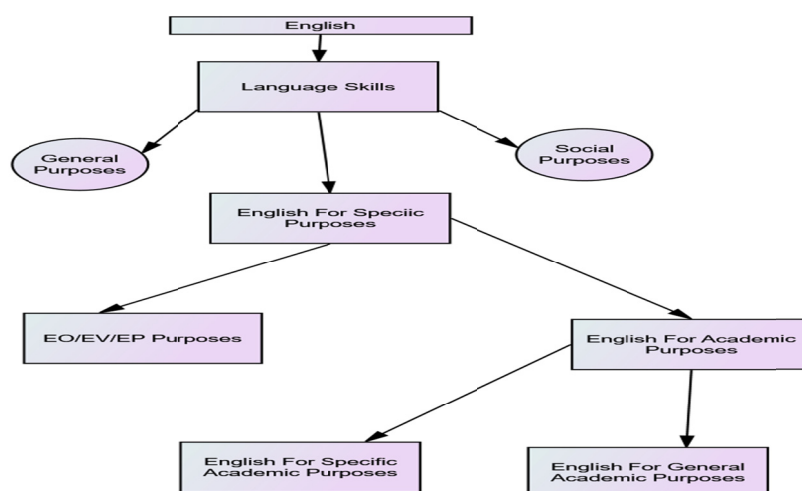


Figure 1. ESP categories tree diagram (Adopted and adapted from Jordan 1997, p. 3)

### 3. Phraseology

#### 3.1 Specialised Corpus

This is a type of corpus designed to represent a particular sort of text used in a specialised discourse such as, law, newspaper editorial, business, technology, geography, academic articles and so forth. Specialised corpus is used as representative of a given type of text and “There is no limit to the degree of specialisation involved” (Hunston, 2002, p. 15). The restriction is often draws based on the time frame, social setting or discourse, subjects and forms language used written or spoken. The notion of specialised is embodied in the studies described in Ghadessy et al. (2001) small corpus in ELT, the term specialised corpora does not necessarily imply that corpora in question are small. Some specialised corpora run into millions of words such as: Cambridge and Nottingham Corpus (CANCODE), and the Michigan Corpus of American Spoken English (MICASE).

Table 1. Parameters for defining corpora as specialised

Parameters	Details/ Examples
Specific purpose for compilation:	To investigate particular grammatical, lexical, lexicogrammatical, discoursal or rhetorical features
Contextualisation:	Setting (e.g. court, school)
Size:	
Whole corpus	1-5 million words
Sub-corpus or small-scale corpus	20, 000-250,000 words
Genre:	Commercial law
Types of text / discourse:	Law textbooks, casual conversations
Subject matter / topic:	Law, Economics, the weather
Variety of English	Learner, non-standard (e.g. Nigerian, Malaysia)

#### 3.2 Phraseology in Specialised Corpus-Based Studies

More recently linguists renew the interest in phraseology triggered through the corpus-based studies; this is a transition in linguistic studies that shift the attention of linguists especially terminologists from word to pattern. Moreover as Biel (2014, p. 176) construes that corpus linguistics studies have revolutionised studies on terminology from the early conception of the dominant approach of Wuster’s “General Theory Terminology” (GTT) to a specialised phraseology which “incorporate social, cognitive and communicative aspects” associated with terms. In addition, the “traditional approach to phraseology with idioms as the prototype was replaced by the frequency-based approach comprising lexicogrammatical patterns which would be regarded as free or peripheral within the traditional approach but recurrence threshold” (p. 176) in the frequency-based approach phraseology is a broader category which includes collocations, multi-word lexical units, and lexical bundles as

Biber et al. (2007, pp. 58-59) define it as “sequences of words that tend to co-occur, irrespective of their idiomaticity and whether or not the sequence of words constitutes a grammatical unit”. However, the study of the terms was maintained and “remained unchanged” the main object of terminology, although it was extended and spilled over to account for phraseology particularly through the collocations of terms.

Corpus linguistics studies present that words in the naturally occurring language are produced and formed together in a vacuum. As Cheng (2012) asserts that “If words collocate, this means that they are co-selected by the speaker or writer and they are not a chance co-occurrence” (p. 77). The notion of collocation supports the specialised use of language in a discourse and that the collocation is the linguistic vehicle that projects the words as used in naturally occurring discourse. Traditionally terminologists study language in terms of specified registers used in different discourse, while corpus linguists imply collocation as an avenue to vertically study the language used in terms of the words (registers) and their the natural co-selections. Linguists at the realm of specialised collocations look at how the word is significantly collocated with others, and why it avoids occurring with others. Studies in specialised collocation reveal that words collocate with other words not by mere chance but are most frequently co-selected in the vicinity of that word.

#### 4. Legal Phraseology

##### 4.1 Legal Language (*Legalese*)

The language of the law as it is called “legalese” has very distinctive lexical and structural patterns which in many ways differ from the “traditional forms of language”. Its conservatism is linked directly to the need for unambiguous language that has already been tried and tested in the courts. By retaining to traditional lexis and structure lawyers can be confident that the language of the law is consistent and precise.

According to Chroma (2011), the language of the law is classified in various ways and one may speak of a language for special or specific purposes, or of a sublanguage, scientific language, or specialized language (Pearson 1998). Widdowson (1979, p. 24) characterizes such subcategories of a language “not as formally defined varieties of English, but as realizations of universal sets of concepts and methods or procedures which define disciplines or areas of enquiry independently of any particular language”; Sinclair (2007) refers this to local grammar describing the specificity of the language in a particular subject area, and what Jackson (1997) terms legal grammar when analysing the semiotics of legislative texts.

According to Ruusila and Lindroos (2016) the term phraseology from the generally perspective, is used describe to “formulaic language”, “fixed expressions ranging from single word combinations such as idioms, collocations and binomials to larger linguistic units such as routine formulae”. Phraseology was part and percept of legal text. And how phraseological units are used across legal genres poses a lot of communicative challenges in legal discourse.

Biel (2014) postulates that based on traditional perspective legalese is viewed as the property of terminology rather than phraseology therefore phraseology treated as “perfunctorily” and not prioritised. In addition, Chroma (2011) posits that legal English has been subject to a more consistent investigation and research since 1963 when Professor Mellinkoff published his book (1963, 2004) with quite a challenging aim to rationalise the language of the law in order to make it more comprehensible not only to the lay audience (e.g. clients) but also to avoid unnecessary ambiguity causing misunderstanding even between lawyers.

Moreover, in the recent studies of legalese using corpus methodologies approximately in the last three decades, the interest was shifted from terminological-based approach to phraseological-based approach. The pioneers of this new paradigm include Kjaer (1990; 2007) on, Gozdz-Rozkowsk (2011) who studies on “multi-word patterns across legal genres” Pontrandolfo (2011) who studies on phraseology in translated criminal judgments, and more recently Biel (2014; 2015) who studies on “phraseology in Legal translation”, Yunus et al. (2016) who study on “colligations of prepositions: essential properties of legal phraseology” and Ruusila and Lindroos (2016). Residing upon two theoretical background “systemic linguistics” and intertextuality theory Kjaer (1990) proposes a classification of norm-conditioned legal word combination, this comprises four items:

- 1) prefabricated word combinations directly prescribed by law.
- 2) word combination only indirectly prescribed by law.
- 3) word combination based on implicit quotation from other texts in a genre chain in the legal domain.
- 4) habitual used routine phrases

However, she postulates that her classification was based on the correlation between the stability of combinations and legal constraints which emphasises nature of the systemic nature of legal phraseology and

intersexuality of the legal texts, showing the link between law and lower-ranking texts in the hierarchy of legal genres: coherence of the system is based on the intertextual relations between legal texts, obtained by reproduction (implicit quotation) and recontextualisation of words and phrases (Biel, 2015).

Campaigners for plain English—*English that is straight forward and easy to understand*—argue that legal language could be simplified so that is both more comprehensible to ordinary people and more practical for lawyers themselves. Many lawyers fear, however, that if changes were made, new simplified language structures could create legal loopholes (Thorne, 2008). Legalese is used in various legal contexts in both written and spoken forms. It is characterised by its “complex grammatical structures”, “technical lexis”, “archaic expression”, and “limited punctuation”, these all together are what “make it quite different from other varieties”. “I swear by Almighty God to tell the truth, the whole truth and nothing but the truth.....”

Furthermore, legal language is the domain of specialist occupation and the intended audience comprises people who are experts in the field. Legalese is recurrently and explicitly used by the specialists in the field, meanwhile it complex formulaic in nature, “it is quite unlike informal speech with its irregular patterns, and quite unlike the language of literature with its personal, often idiosyncratic approach, instead it draws on structures that have been predefined and pretested, and uses jargon that is familiar only to the experts and grammatical structures that are difficult to decode” (Thorne, 2008). The distinctive lexis and syntax elements of legalese are what make it difficult to comprehend by the laymen.

## 5. The Distinctive Elements of Legal English

### 5.1 The Elements of Legal Texts and Their Characteristics

To grasp the peculiarities in legal thinking, and to understand the legal constructs behind the terms and phrasemes used in legal language, linguists especially terminologists and phrenologists use various elements at the linguistic surface of legal language (legalese). These consist of the use of “fixed word patterns”, “routine expressions” and “formulaic use of prefabricated terms” (Randall & Graf, 2014) to achieve precision and persuasion in a certain oral communicative situations in legal discourses like court proceedings and in diverse types of written legal texts such as contracts, legislative texts, judgments, and powers of attorney. Legal texts have distinctive linguistic feature pertaining to style, i.e. the choice and positioning of language elements to express legal substance in legal texts.

Passive verbs (Gough 1966; Randall & Graf, 2014), nominalizations (Klare & Smart, 1973), negatives (Cutler, 1983), omitted arguments; low-frequency, formal-register, and presupposed or undefined terminology (Randall & Graf 2014), such instructions have been shown to confuse even highly educated jurors, never mind jurors with less education or non-native English skills (Charrow & Charrow 1979; Elwork, et. al. 1982; Tiersma, 1999).

It should also be noted that legal language clusters may be very long, ranging from phrases, sentences, to entire clauses or parts of documents, known as boilerplate clause (Tierma, 1999; Biel, 2014). Research into phraseology in legal documents can help to shed light on the frequency and nature of the phraseological patterns used to construct legal texts and, through the analysis of recurring word combinations, can help to reveal linguistically expressed thought patterns that are rooted in the history of a legal system and embedded in a given legal culture.

### 5.2 Linguistic Inquiries in Lexical Semantics

Semantics is the “study of the linguistic meaning of morphemes, words, phrases, and sentences”. Lexical semantics, being a subfield of semantics, is about “the meanings of words and the meaning relationships among words” (p. 584). Halliday (1993) discusses the meaning relationship between word forms and the grammatical contexts in which they occur; for instance, *grow* (verb) and *growth* (noun) construe different realities not due to their different morphology but due to the different environments in which they occur.

Halliday and Hassan (2014) construes that words are the obvious elements of a language, in our speech we utter words and when we write on the page we produce words. But by looking more closely to a word on the page, a word may denote different meaning far from the simple and obvious dictionary definition we imagine it to be. He concludes that “the borderline between content words and function words is not a sharp one: rather, the two form a continuum or cline, and words like always and top lie somewhere along the middle of the cline. Thus there is no exact point where the lexicologist stops and grammarian takes over; each one can readily enter into the territory of the other” (p. 10).

### 5.3 The Semantic Functions of Linguistic Elements in Legal Contexts

Many Semanticists distinguish between word as an independent unit for paradigmatic and morphological

purposes and lexical unit having not only a particular form but also a particular sense (Chroma, 2011; Halliday 2004; Hoey, 2005). A word in its narrower sense can hardly be used as a unit of meaning but only when uses in its wider sense and interchangeably with the term “lexical unit”. What will be strictly distinguished, following Löbner, 2002, is a concept for mental description or meaning and term for its spelling (or sound) form. However, there are different types of linguistic elements this study is limited to prepositions.

## 6. Conclusion

This paper presented issues concerning phraseology in language for specific legal purposes. It gave some insights on the application of corpus-based approach to studies of legal phraseology. From this overview we can draw three implications: theoretical, methodological, and research implications. Theoretically, the paper presented some of the sheer issues related to English for specific purpose and phraseology in specific legal purpose. This adds to the literature, in this field. Methodologically, it gave some insights on studies of phraseology using corpora and concordance tools as a method of analysis. This will benefit novice researchers' interesting in phraseology and corpus linguistics research. Consequently, law students in tertiary education need to possess knowledge of legal phraseology and legal terminologies. The students when approaching a language with specific scope are required to have a distinctive measure between end and means. The interesting question about the specificity of the language for specific purpose is the used of terms in a specific genre, for instance the technical language uses technical words, scientific language uses scientific words, and the language of the particle physics uses particular physics words, and so on.

## References

- Alexander, R. J. (1984). Fixed expressions in English: reference books and the teacher. *ELT Journal*, 38(2), 127-134. <https://doi.org/10.1093/elt/38.2.127>
- Bardovi-Harlig, K. (2002). A new starting point?: Investigating formulaic use and input in future expression. *Studies in second language acquisition*, 24(2), 189-198. <https://doi.org/10.1017/S0272263102002036>
- Bathia, V., Hernandez, P. S., & Pérez-Paredes, P. (2011). Specialized languages Corpora, meta-analyses and applications. *Researching Specialized Languages*, 47, 1. <https://doi.org/10.1075/scl.47.02bha>
- Biber, D., Conrad, S., & Cortes, V. (2004). If you look at...: Lexical bundles in university teaching and textbooks. *Applied linguistics*, 25(3), 371-405. <https://doi.org/10.1093/applin/25.3.371>
- Biel, Ł. (2015). Phraseological profiles of legislative genres: complex prepositions as a special case of legal phrasemes in EU law and national law. *Fachsprache*, 139-161. <https://doi.org/10.24989/fs.v37i3-4.1286>
- Bloor, T., & Bloor, M. (1995). *The functional analysis of English*. London, New York: Arnold.
- Bobrow, S. A., & Bell, S. M. (1973). On catching on to idiomatic expressions. *Memory & Cognition*, 1(3), 343-346. <https://doi.org/10.3758/BF03198118>
- Boers, F., Demecheleer, M., He, L., Deconinck, J., Stengers, H., & Eyckmans, J. (2017). Typographic enhancement of multiword units in second language text. *International Journal of Applied Linguistics*, 27(2), 448-469. <https://doi.org/10.1111/ijal.12141>
- Cacciari, C., & Glucksberg, S. (1991). Understanding idiomatic expressions: The contribution of word meanings. In *Advances in psychology* (Vol. 77, pp. 217-240). North-Holland. [https://doi.org/10.1016/S0166-4115\(08\)61535-6](https://doi.org/10.1016/S0166-4115(08)61535-6)
- Charrow, R. P., & Charrow, V. R. (1979). Making legal language understandable: A psycholinguistic study of jury instructions. *Columbia law review*, 79(7), 1306-1374. <https://doi.org/10.2307/1121842>
- Chomsky, N. (1981). A note on non-control PRO. *Journal of Linguistic Research*, 1(4), 1-11.
- Chomsky, N. (1986). *Knowledge of language: Its nature, origin, and use*. Greenwood Publishing Group.
- Chromá, M. (2011). Synonymy and polysemy in legal terminology and their applications to bilingual and bilingualtranslation. *Research in Language*, 9(1), 31-50. <https://doi.org/10.2478/v10015-011-0004-2>
- Cowie, A. P. (2002). *English dictionaries for foreign learners: A history*. Oxford University Press on Demand.
- Cutler, A. (1983). Lexical complexity and sentence processing. In *The process of language understanding* (pp. 43-79). Wiley.
- Dudley-Evans, T. (1997). Five questions for LSP teacher training. *Teacher education for LSP*, 58-67.
- Dudley-Evans, T., & St John, M. J. (2012). *Developments in English for specific purposes: A multi-disciplinary approach*. Cambridge University Press.

- Elwork, A., Sales, B. D., & Alfini, J. J. (1982). *Making jury instructions understandable*. Michie Company.
- Firth, J. R. (1968). *Selected papers of J.R. Firth, F.R. Palmer*. London, UK.
- Flowerdew, J. (2013). *Discourse in English language education*. New York: Routledge.
- Fromkin, V., Rodman, R., & Hyams, N. (2018). *An introduction to language*. Cengage Learning.
- Ghadessy, M., Henry, A., Roseberry, R. L. (Eds.). (2001). *Small Corpus Studies and ELT: Theory and Practice*. Amsterdam, Philadelphia: John Benjamins Publishing Company. <https://doi.org/10.1075/scl.5>
- Gollin-Kies, S., Hall, D. R., & Moore, S. H. (2016). *Language for specific purposes*. Springer.
- Gough, P. B. (1966). The verification of sentences: The effects of delay of evidence and sentence length. *Journal of Verbal Learning and Verbal Behavior*, 5(5), 492-496. [https://doi.org/10.1016/S0022-5371\(66\)80067-1](https://doi.org/10.1016/S0022-5371(66)80067-1)
- Halliday, M. A. (1993). Towards a language-based theory of learning. *Linguistics and education*, 5(2), 93-116. [https://doi.org/10.1016/0898-5898\(93\)90026-7](https://doi.org/10.1016/0898-5898(93)90026-7)
- Halliday, M. A. K., Matthiessen, C., & Halliday, M. (2014). *An introduction to functional grammar*. Routledge. <https://doi.org/10.4324/9780203783771>
- Hasselgard, H. et al. (2013). *Corpus perspectives on patterns of lexis: studies in corpus linguistics*. Amsterdam, Philadelphia: John Benjamins Publishing Company. <https://doi.org/10.1075/scl.57>
- Hoey, M. (2005). *Lexical priming: A new theory of words and language*. Routledge. <https://doi.org/10.4324/9780203327630>
- Hunston, S. (2002). *Corpora in applied linguistics*. Ernst Klett Sprachen. <https://doi.org/10.1017/CBO9781139524773>
- Hyland, K. (2006). *English for academic purposes: An advanced resource book*. Routledge.
- Jordan, R. R. (1997). *English for academic purposes: A guide and resource book for teachers*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511733062>
- Klare, G. R., & Smart, K. L. (1973). Research Briefs. *Journal of Educational Research*, 67(4), 176. <https://doi.org/10.1080/00220671.1973.10884594>
- Löbner, S. (2002). *Understanding Semantics*. London: Arnold, Hodder Headline Group.
- Mitkov, R. (Ed.). (2017). *Computational and Corpus-based Phraseology: Second International Conference, Europhras 2017, London, Uk, November 13-14, 2017, Proceedings* (Vol. 10596). Springer.
- Naciscione, A. (2017). *Stylistic use of phraseological units in discourse*. John Benjamins Publishing Company.
- Ohlrogge, A. (2009). Formulaic expressions in intermediate EFL writing assessment. *Formulaic language*, 2, 375-386. <https://doi.org/10.1075/tsl.83.07ohl>
- Paltridge, B. (2013). Genre and English for specific purposes. *The handbook of English for specific purposes*, 347-366.
- Partington, A. (1998). *Patterns and meanings: Using corpora for English language research and teaching*. John Benjamins Publishing. <https://doi.org/10.1075/scl.2>
- Partington, A. (2004). Utterly content in each other's company: Semantic prosody and semantic preference. *International journal of corpus linguistics*, 9(1), 131-156. <https://doi.org/10.1075/ijcl.9.1.07par>
- Pawley, A., & Syder, F. H. (1983). Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. *Language and communication*, 191, 225.
- Philip, G. (2011). *Colouring meaning: Collocation and connotation in figurative language*. John Benjamins Publishing. <https://doi.org/10.1075/scl.43>
- Rammell, C., Van Lancker, S., & Pisoni, D. B. (2018). Perception of formulaic and novel expressions under acoustic degradation. *The Mental Lexicon*, 12(2), 234-262. <https://doi.org/10.1075/ml.16019.ram>
- Randall, J. H., & Graf, L. R. (2014). Linguistics meets "legalese": syntax, semantics, and jury instruction reform. In *LSA Annual Meeting Extended Abstracts*, 5(1). <https://doi.org/10.3765/exabs.v0i0.2381>
- Rashid, R. A. (2016). Topic continuation strategies employed by teachers in managing supportive conversations on Facebook Timeline. *Discourse Studies*, 18(2), 188-203. <https://doi.org/10.1177/1461445615623906>
- Rashid, R. A., Yunus, K., & Wahab, Z. (2018). *Supportive conversations on Facebook Timelines: Discourse*



- topic management*. Abingdon, Oxford: Routledge. <https://doi.org/10.4324/9781351057714>
- Shaikhullin, T., & Alavi, A. (2015). Linguistic representation of family relations in russian and yemeni set expressions.
- Sinclair, J. (1991). *Corpus, Concordance, Collocation: Describing English Language*. Oxford: Oxford University Press.
- Sinclair, J. (2004). *Trust the text: Language, corpus and discourse*. New York: Routledge.
- Strecker, B. (1985). Rules and the dynamics of language.
- Stubbs, M. (1995). Collocations and semantic profiles: On the cause of the trouble with quantitative studies. *Functions of language*, 2(1), 23-55. <https://doi.org/10.1075/fo1.2.1.03stu>
- Stubbs, M. (2001) Texts, corpora, and problems of interpretation: a response to Widdowson. *Applied Linguistics*, 22, 149-172. <https://doi.org/10.1093/applin/22.2.149>
- Thorne, S. (1998). *Mastering advanced English language*. New York: Macmillan.
- Tiersma, P. M. (1999). *Legal language*. University of Chicago Press.
- Titone, D. A., & Connine, C. M. (1999). On the compositional and noncompositional nature of idiomatic expressions. *Journal of pragmatics*, 31(12), 1655-1674. [https://doi.org/10.1016/S0378-2166\(99\)00008-9](https://doi.org/10.1016/S0378-2166(99)00008-9)
- Tognini-Bonelli, E. (2001). *Corpus linguistics at work* (Vol. 6). John Benjamins Publishing, USA. <https://doi.org/10.1075/scl.6>
- Viana, V., Zyngier, S., & Barnbrook. (2011). *Perspectives on corpus linguistics: studies in corpus linguistics*. Amsterdam/Philadelphia: John Benjamins Publishing Company.,
- White, L. (1989). *Universal grammar and second language acquisition* (Vol. 1). John Benjamins Publishing. <https://doi.org/10.1075/lald.1>

### 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/>).