ISSN: ۲۶۷۶-۵۷۶٤
www.jocrisar.ir
مجله پژوهش های معاصر در علوم و تحقیقات
سال هفتم، شماره ۷۷، آذر ۱۴۰۴

Developing an Academic Word List for the Field of Law: A Corpus-Based Investigation

Rouhollah Karami
M.A. in English Language, specializing in Teaching English as a Foreign Language (TEFL),
Faculty of Literature and Humanities,
Ilam University, Iran.

Abstract
This study had two main goals: first, to analyze a corpus of ۱۲۰۰ law research articles and identify the most commonly used words, and second, to compare this list to Coxhead's Academic Word List. Theresearcherused a corpus of ٤,۱۱٤,۳۲۸ words and set a frequency cutoff of ۱۲۰ occurrences. After analyzing the corpus and excluding function words, bibliographies, tables, headers, and footnotes, a list of ۳۷۳۰ academic words for the field of law was created. The resulting word list was then subjected to expert review and validation by a panel of law professionals to ensure that the words on the list are relevant, useful, and appropriate for academic context. ۲٥٤۱ words were deemed unsuitable and consequently retained from the final word list. On the other hand, ۱۱۸۹ words successfully passed the validation process and were approved for inclusion in the finalized academic word list.۳۷۲ word families were identified .Of these, ۱۲۱ word families(۳۲,٥۲٪) overlapped with Coxhead's Academic Word List(AWL), while the remaining ۲٥۱ wordfamilies(٦۷,٤۸٪) were unique to the Law Academic Word List, highlighting the specialized terminology related to law field. This study has important implications for course designers, ESP experts, authors, and students studying law at both the undergraduate and postgraduate levels.

Keywords: Academic Word List, Law Academic Word List , Corpus

Introduction

A discourse community, as Swales (۱۹۹۰, p. ۹) defines, is a "socio-rhetorical network that exists" to attain particular shared goals, and to reach these goals, members of the discourse community require familiarity with the particular inter-communication mechanisms or genres which are characteristic of that community. Central to the notion of discourse community is genre, which is "a class of communicative events, the members of which share some set of communicative purposes" (Swales, ۱۹۹۰, p. ۵۸). Moreover, genres can be described as activities of a discourse community which, according to Swales, are "determinant of linguistic choices (۱۹۹۰, p. ٤۲); therefore, investigating the conventionalized modes of expression of a particular genre can be very revealing. These modes of expression could be spoken or written (among other modes such as electronic or multimodal). Written modes of expression, in particular, academic writing is of paramount significance in academic communities since it is a requirement for participation in higher education (Bruce, ۲۰۰۸, p. ۱), and a medium "through which scholars attempt to negotiate with and persuade other scholars of the legitimacy of their arguments" to be recognized as accepted knowledge within their disciplines (Hewings, ۲۰۰۱, p. ۱۰). Furthermore, of all genres "within academic writing, it is the research article that has attracted the most attention" as the primary medium for presenting claims of new knowledge (Hewings, ۲۰۰۱, p. ۱۲-۱۳). Moreover, "academic discourse refers to the various ways of thinking and using language that are present in the academy" (Hyland, ۲۰۰۹a, p. ۱). The study of academic discourse is a young but swiftly rising field. Academic discourse is gaining more prominence as more students in higher education need to communicate within their academic communities, as scholars' careers are tied to their success in publishing, and as social institutions, media, and advertising are influenced by the language of science more than ever (Hyland, ۲۰۰۹a, p. viii). This empirical language of science based on logic and argument is the basis of scientific progress and the construction of knowledge.As a useful relationship, corpus-based linguistics, which is a "methodological innovation" to study aspects and features of language based on huge textual corpora (Lee, ۲۰۰۸), and academic discourse analysis can be good partners. The accessibility of banks of naturally occurring texts on computers has made it possible to analyze language structure and use it in unprecedented ways to gain new insights (Jalilifar, ۲۰۱٤, p. ۲۷۱). Furthermore, corpus-based linguistics can enhance the scientific progress and career development of both students and academics around the world. In order to do so, corpora need to be designed around academic genres from the ground up (Lee ۲۰۰۸). Arguably, the most influential genre within academic discourse and subsequently English for Academic Purposes (EAP) is the research article (RA) (Hyland, ۲۰۰۹a, pp. ٦۷-٦۸), and corpus-based linguistics can offer researchers tools to analyze academic texts [such as research articles as a single genre] for concordances, keywords, and wordlists (Jalilifar, ۲۰۱٤, p. ۲۷۹).To corpus linguists, the notion of frequency is central, as corpora are concerned with what frequently occurs (Hyland, ۲۰۰٦, p. ۵۸). In other words, the underlying assumption is that if a linguistic feature is observed to occur frequently and typically in the past, then it is probable to be significant in the future as well (p. ۵۹). One use of frequency counts in EAP is the production of academic vocabulary lists such as Coxhead's (۲۰۰۰) Academic Word List (AWL), since, in line with Li

and Pemberton (as cited in Hyland, ۲۰۰٦), students are believed to find academic vocabulary very challenging to learn.

The research article (RA) is the most prominent genre in academic writing as it serves as the primary means of presenting new knowledge (Hewings, ۲۰۰۱, p. ۱۲). Despite the availability of alternative electronic publishing platforms, the research article remains the preeminent genre of academia and the primary site for disciplinary knowledge-making (Hyland, ۲۰۰۹, p. ٦۷). It is considered the "master narrative" of our time (Montgomery, ۱۹۹٦, p. ۲). The significance of the research article can be attributed to the highly esteemed peer review process, which serves as a regulating system for transforming ideas into knowledge (Hyland, ۲۰۰۹, p. ٦۸). Authors aim to have their arguments become part of the disciplinary consensus by contextualizing new findings within previously published literature, effectively displaying the current status of the existing consensus (Hewings, ۲۰۰۱, p. ۱۲). Additionally, the research article holds prestige as a genre that creates a means for scientific fact-generation as more academics worldwide are required to publish in high-impact, peer-reviewed journals for career development and promotion (Hyland, ۲۰۰۹, p. ٦۷), a point also acknowledged by other scholars (e.g., Hartley, ۲۰۰۸; Swales, ۱۹۹۰).

Several studies have been conducted to create word lists that represent the most common academic lexical items due to the significance of academic vocabulary. Two approaches have been adopted for this purpose. The first approach is the common core approach, which aims to provide discipline-crossing academic word lists appropriate for learners in various fields (Campion & Elley, ۱۹۷۱; Coxhead, ۲۰۰۰;Ghadessy, ۱۹۷۹; Lynn, ۱۹۷۳; Praninskas, ۱۹۷۲). The second approach is the early specialization approach, which intends to create discipline-specific academic word lists based on independent domains suitable for the learners of that particular field. Xu and Nation (۱۹۸٤) produced a University Word List (UWL) based on the lists developed by Campion and Elley (۱۹۷۱), Ghadessy (۱۹۷۹), Lynn (۱۹۷۳), and Praninskas (۱۹۷۲), which was extensively used for ۱۵ years. However, in an attempt to construct a more robust and more representative academic word list, Coxhead (۲۰۰۰) introduced her Academic Word List (AWL) in a seminal article. The AWL (Coxhead, ۲۰۰۰) consists mainly of academic words and accounts for roughly ۱۰,۰٪ of all the running words in academic texts (Coxhead's corpus) but merely ۱,٤٪ of all the running words in a fiction corpus of the same size.

Literature Review

Academic word lists

The importance of academic vocabulary has led to numerous studies aimed at creating word lists representing the most common academic lexical items. Two approaches are employed: the common core approach, favored by early researchers, focuses on creating discipline-crossing academic word lists suitable for learners in various fields; while the early specialization approach aims to develop discipline-specific academic word lists based on independent domains, tailored for learners in that specific field.The Academic Word List (AWL), developed by Coxhead (۲۰۰۰), comprises ۵۷۰ word families, organized according to the Bauer and Nation scale (۱۹۹۳) up to level ٦. The corpus used for its creation, containing

۳,۵ million tokens, was divided into four sections: arts, commerce, law, and science, each with approximately ۸۷۵,۰۰۰ tokens. Coxhead (۲۰۰۰) utilized ٤١٤ texts, balanced for length and drawn from various sources, including textbooks, articles, book chapters, and manuals. The goal was to produce a vocabulary list beneficial to any L۲ English student in an English-medium academic environment.The AWL was designed as a supplement to the General Service List (GSL), excluding any words from the GSL. Word inclusion was based on frequency, range, and uniformity: each word had to occur at least ۱۰۰ times across the corpora (frequency), in at least ۱۵ of the ۲۸ subject areas (range), and over ۱۰ times in each of the four sub-corpora (uniformity). Coxhead also ensured text balance in length and avoided author bias, representing ٤۰۰ separate authors in her ٤١٤ texts. The AWL's purpose, according to Coxhead, was to assist EAP teachers in selecting vocabulary for their students to learn (Coxhead, ۲۰۱۱).The AWL has been widely used in ESP research, achieving an average coverage of approximately ۱۰٪ across various studies (e.g., Cobb & Horst, ۲۰۰٤; Li & Qian, ۲۰۱۰; Martinez, Beck & Panza, ۲۰۰۹). The AWL remains relevant to modern vocabulary studies and is frequently utilized in the creation of specialist word lists (e.g., Coxhead & Hirsh ۲۰۰۷; Konstantakis, ۲۰۰۷, ۲۰۱۰; Wang et al., ۲۰۰۸).The New Academic Vocabulary List (AVL)Gardner and Davies (۲۰۱۳) developed a new Academic Vocabulary List (AVL) based on a ۱۲۰-million-word academic subcorpus of the ٤۲٥-million-word Corpus of Contemporary American English (COCA). They used lemmas, rather than word families, for initial counting and analysis, later forming word families to address specific academic needs.Gardner and Davies (۲۰۱۳) noted the growing prominence of the Academic Word List (AWL) in education (e.g., Hiebert and Lubliner ۲۰۰۸; Baumann and Graves ۲۰۱۰; Nagy and Townsend ۲۰۱۲), while acknowledging concerns about the widening achievement gap. They identified two limitations of the AWL: the use of word families to determine word frequencies and its relationship to the General Service List (GSL).In their research (۲۰۱۳), Gardner and Davies employed four criteria to define core academic vocabulary: ratio (word frequency must be at least ٥۰٪ higher in the academic corpus than in the non-academic portion of COCA), range (present in at least seven of the nine academic disciplines), dispersion (at least ۰,۸۰), and discipline measure (word frequency cannot exceed three times the expected frequency per million words in any of the nine disciplines). The corpus encompasses nine disciplines: education, humanities, history, social science, philosophy, law and political science, science and technology, medicine and health, and business and finance, with ۸٥ million of the ۱۲۰ million words originating from academic journals.Gardner and Davies (۲۰۱۳) concluded that the AVL covers ١٤٪ of academic materials in both COCA and the British National Corpus (BNC).Discipline-crossing academic word lists

Specialized Academic Word Lists

Recent research in academic vocabulary has emphasized the need for discipline-specific lists of essential vocabulary (Hyland & Tse, ۲۰۰۷). This has led to the creation of lists that either slightly modify the Academic Word List (AWL) (e.g., for medicine: Chen & Ge, ۲۰۰۷; for agriculture: Martinez et al., ۲۰۰۹; for applied linguistics: Vongpumivitch, et.al, ۲۰۰۹) or develop entirely new discipline-specific lists (e.g., for medicine: Wang, Liang, & Ge, ۲۰۰۸). However, some researchers have questioned the representativeness of these lists, citing

limited corpus size and disciplinary distribution (e.g., four separate studies).Several studies have investigated the coverage of the AWL in specific fields, given its status as the most commonly used academic words in academic texts. For example, Vongpumivitch, Huang & Chang (۲۰۰۹) analyzed a ۱,۵ million word corpus of ۲۰۰ research articles in applied linguistics and found that ۱۱,۱۷٪ of the words were from the AWL. They also identified ٤٧٥ AWL word forms and ۱۲۸ non-AWL word forms that appeared frequently in the corpus, suggesting the potential benefits of field-specific academic word lists for better understanding published academic texts. Chen and Ge (۲۰۰۷) examined the frequency and distribution of Coxhead's AWL word families in medical research articles. They found that the AWL accounted for approximately ۱۰,۰۷٪ of the running words, a reasonable proportion. However, only ٥۱,۲٪ of the AWL words appeared with any frequency in medical research articles, while low-frequency words in these articles were frequently found in the AWL. Their corpus consisted of ٥۰ medical research articles with ۱۹۰,٤۲٥ running words. Martinez et al. (۲۰۰۹) investigated the advantages and limitations of Coxhead's (۲۰۰۰) AWL in agriculture research articles. They highlighted the list's value in constructing scientific experience, its coverage in academic texts (around ۱۰٪), and its contribution to non-native English speakers' understanding. However, they also noted that the AWL words were not evenly distributed across specialized academic texts. Additionally, they pointed out that the meaning and use of AWL words can vary across disciplines, emphasizing the need for field-specific and discipline-based academic word lists that incorporate all frequent academic lexis items necessary for expressing the specific research area's rhetoric.Hyland and Tse (۲۰۰۷) conducted a study to analyze the distribution and frequency of AWL words in a ۳,۳ million word corpus of research articles, textbook chapters, academic book reviews, and scientific letters. Their goal was to determine whether Coxhead's AWL vocabulary had the same value across different disciplines and fields. Their corpus-based evidence demonstrated that AWL items vary significantly across disciplines and fields. The same word can exhibit differences in frequency, range, preferred meanings and forms, and collocation patterns (Hyland and Tse, ۲۰۰۷, p.۲۳۸). While the AWL covered a similar proportion of their total corpus as proposed by Coxhead, the coverage was not consistent across all disciplines. Although the combined General Service List (GSL) and AWL covered approximately ۸٥٪ of the overall corpus, there were differences in coverage when the distribution was analyzed by sub-fields. Researchers have recognized that the AWL, while valuable, does not contain words of equal importance for all students, and some words may be irrelevant to specific disciplines. This has led to a call for more focused, discipline-specific word lists (e.g., Li & Qian, ۲۰۱۰). Li and Qian (۲۰۱۰) analyzed a ۳,۳ million-word corpus of financial texts, finding that the AWL covered ۱۰,٤٦٪ of the words. They also noted the significant presence of high-frequency AWL items, with a cumulative coverage of ۲۲,۰۳٪. This emphasis on discipline-specific lists has been echoed by other researchers. Lam (۲۰۰۱), for instance, studied academic words used in computer science, finding that the same words can have different meanings in academic contexts compared to general texts. For example, students might understand "field" in an agricultural sense but not as a "data field" in a database program. Lam proposed creating glossaries of academic vocabulary with frequency information based on specialized corpora.

Mudraya (۲۰۰۶) developed a ۱,۲۰۰-word family list for engineering students, representing vocabulary likely encountered in engineering textbooks. Mudraya's study highlighted the importance of sub-technical vocabulary and academic English in ESP classrooms.Wang, Liang, and Ge (۲۰۰۸) created a Medical Academic Word List of ٦٢٣ word families, accounting for ۱۲,۲٤٪ of the tokens in their corpus of ۱,۰۹۳,۰۱۱ running words from medical research articles. They also advocated for discipline-specific word lists.Ward (۲۰۰۹) developed a word list for engineering students based on a ۲۷۱,۰۰۰-token corpus, arguing that it would be more useful for students across disciplines due to its coverage, frequency, and manageable size. Ward identified ۲۹۹ frequent items, including non-technical words that "represent a relatively easy target for learners whose high school education has not equipped them for the linguistic challenges they face in reading English language textbooks" (p. ۱۸۰).

Recent studies have shed light on the specific vocabulary demands of different academic fields. Valipouri and Nassaji (۲۰۱۳) investigated academic words used in chemistry, analyzing a corpus of ٤ million words from ۱,۱۸٥ chemistry research articles (RAs) across four main subject areas. They found that the Academic Word List (AWL) covered ۹,٦۰٪ of their corpus. Similarly, Yang (۲۰۱٥) explored frequently used nursing vocabulary, creating a word list of ٦٧٦ word families from ۲٥۲ nursing RAs, representing approximately ۱۳,٦٤٪ of the corpus.This extensive research has significantly advanced our understanding of vocabulary distribution across various academic domains. It has become clear that the lexical challenge in academic writing is even greater than previously thought. This improved understanding is largely due to corpus-based vocabulary research.Pioneering work by West (۱۹٥۳) and more recent efforts by Coxhead (۲۰۰۰) have significantly contributed to focusing research on academic vocabulary. However, research by Hyland and Tse (۲۰۰۷) and subsequent studies on disciplinary vocabulary (e.g., Martinez et al., ۲۰۰۹; Wang, Liang, & Ge, ۲۰۰۸) have highlighted the need for even more focused research. This research suggests significant cross-disciplinary variation in lexical use and distribution. Furthermore, we know that different disciplines employ varying levels of lexical diversity, with some disciplines using more words than others (Biber, ۲۰۰٦). This disciplinary variation raises questions about our understanding of lexical distinctions in academic writing and should temper the strength of conclusions drawn from academic corpora.

Results

This chapter presents the findings of a corpus-based study that aimed to create an English Law Academic Word List (LAWL) by analyzing vocabulary used in law research papers. The chapter is organized into three sections:

۱.The results obtained from Antwordprofiler.

۲.The law field most frequent words and their frequencies

۳.The results of validation process.

Details of the corpus

The Law Research Article Corpus (LRAC) information is shown in table ۱.

Table ۱ The information of Law Research Article Corpus (LRAC)

Corpus Size	**٤,١١٤,٣٢٨ running words**
Extract Text	Full academic papers with the removal of diagrams, graphs, charts, bibliographies, text headers, footnotes, author's name and affiliate, generic names, family names and other parts of the texts which cannot be processed by concordance program
Number of Texts	١٢٠٠ papers
Medium	Written materials in word format
Subject	Law
Language	Papers written in English (Monolingual Corpus)
Publication date	During the period January ٢٠١٥-٢٠٢٣ in order to avoid outdated words in the corpus

This table provides comprehensive details about the nature, scope, and content of the corpus used for the creation of an academic word list within the context of law, including the size, content, format, subject, language, and publication timeframe.

Law academic word list coverage

In this section, we will showcase the results of our data analysis and compare them to a well-established benchmark list. The distribution of AWL and non-AWL words in our corpus (LAWL) is presented in Table ٤,٢, both in terms of word families and word tokens. The AWL was created by Averil Coxhead, a linguist and researcher from New Zealand, to identify words commonly used in academic texts but not in everyday language. It is widely used in EAP programs and educational materials to help students expand their academic vocabulary and improve their proficiency in comprehending and producing academic content.After analyzing the data using "Antwordprofiler," the software generated the results in approximately ١٠ minutes, which are displayed in an accompanying image.

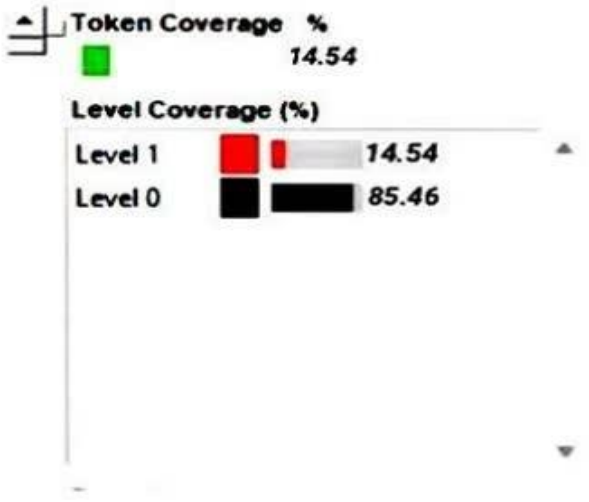

Figure ۱. Software's generated results

The software also generates a detailed table that presents comprehensive information on coverage details, ensuring accuracy and facilitating a more in-depth understanding of the data. This table serves as a valuable resource for users to analyze and interpret the information effectively, enhancing their ability to make informed decisions.

The visual representation shown below is the result of using the "Coverage Stats" feature. It's crucial to recognize that there is a cutoff at ۱۲۰, indicating that a significant number of headwords would usually be excluded based on this criterion. This detail emphasizes the significance of taking the established cutoff into account when interpreting the data, as it has a significant impact on whether headwords are included or excluded in the analysis.

Statistics

LEVEL	FILE	TOKEN	TOKEN%	CUMTOKEN%	TYPE	TYPE%	CUMTYPE%	GROUP	GROUP%	CUMGROUP%
1	3_awl_570.txt	598096	14.54	14.54	2778	2.81	2.81	569	0.59	0.59
0	-	3516232	85.46	100	96067	97.19	100	96067	99.41	100
TOTAL:	4114328			98845			96636			

Figure ۲. Word list coverage details

Table ۲ Analysis of Word Tokens in the Law Field Corpus

Category	Number of Tokens	Percentage of Total Tokens
Total Word Tokens	٤١١٤٣٢٨	١٠٠
Overlap Word Tokens (Coxhead's AWL)	٥٩٨٠٩٦	١٤,٥٤
Unique Word Tokens (Law Field)	٣٥١٦٢٣٢	٨٥,٤٦

Table ٤,٢ provides an overview of the word token distribution in the Law Field Corpus, totaling ٤,١١٤,٣٢٨ tokens. Of these, ١٤,٥٤٪ (٥٩٨,٠٩٦ tokens) overlap with Coxhead's Academic Word List (AWL), representing general academic vocabulary commonly used across disciplines. The remaining ٨٥,٤٦٪ (٣,٥١٦,٢٣٢ tokens) are unique to the law field, reflecting the specialized nature of legal language. This highlights the significance of both academic and technical vocabulary in legal texts, suggesting that proficiency in law requires mastery of both general academic terms and field-specific terminology. These findings support Paul Nation's model of vocabulary classification and emphasize the need for targeted vocabulary instruction in legal education.

Table ٣ Analysis of Word Types in the Law Field Corpus

Category	Number of Types	Percentage of Total Types
Total Word Types	٩٨٨٤٥	١٠٠
Overlap Word Types(Coxhead's AWL)	٢٧٧٨	٢,٨١
Unique Word Types (Law Field)	٩٦٠٦٧	٩٧,١٩

Table ٣ examines the distribution of word types in the Law Field Corpus, which includes a total of ٩٨,٨٤٥ unique word types . Just ٢,٨١٪ (٢,٧٧٨ types) are part of Coxhead's Academic Word List (AWL), representing general academic vocabulary. Conversely, ٩٧,١٩٪ (٩٦,٠٦٧ types) are distinct to the law domain, illustrating the dominance of specialized terminology in legal texts. This analysis reveals the crucial role of technical vocabulary in the law field and suggests that proficiency requires familiarity with both academic and field-specific terms.

Table ٤ Distribution and Validation of Word Types and Tokens in the Corpus

Description	Value	Percentage	Reference
Total word types in the corpus	٩٨٨٤٥	١٠٠	Entire corpus
Word types occurring >١٢٠ times	٣٧٣٠	٣,٧٧	Of total word types
Word types excluded during validation process	٢٥٤١	٦٨,١٢	Of words occurring >١٢٠ times
Word types retained after validation process	١١٨٩	٣١,٨٨	Of words occurring >١٢٠ times
Word families identified	٣٧٢	-	From retained word types

Total word tokens of validated word types	۹۷۵۵۱۷	-	From retained word types

Table ٤ provides an overview of the distribution and validation process of word types and tokens within the corpus. The total number of word types in the corpus is ۹۸,۸٤٥, with ۳,۷۳۰ word types occurring more than ۱۲۰ times, representing ۳,۷۷٪ of all word types. During the validation process, ۲,٥٤۱ word types (or ٦۸,۱۲٪ of those occurring >۱۲۰ times) were excluded, leaving ۱,۱۸٥ word types retained, which account for ۳۱,۸۸٪ of the high-frequency words. From these retained word types, ۳۷۲ word families were identified. The total number of word tokens associated with the validated word types is ۹۷٥٥۱۷. This analysis demonstrates the rigorous validation process applied to ensure the relevance and significance of the selected word types, highlighting their importance in understanding the core vocabulary of the corpus. These findings support the reliability of the data and its applicability to further linguistic analysis.

Table ٥ Overlap and Uniqueness of Law Field Academic Word List (LAWL) with Coxhead's Academic Word List(AWL)

Metric	Value	Percentage	Details
Number of LAWL word families	۳۷۲	۱۰۰	Entire Law field Academic Word List
Number of LAWL overlapping word families with AWL	۱۲۱	۳۲,٥۲	Overlap with Coxhead's AWL
Number of word families unique to LAWL	۲٥۱	٦۷,٤۸	Unique to LAWL, not in AWL

Table ٥ compares the Law Field Academic Word List (LAWL) with Coxhead's Academic Word List (AWL). The LAWL comprises ۳۷۲ word families, of which ۱۲۱ (۳۲,٥۲٪) are shared with the AWL, representing general academic vocabulary used in various disciplines. On the other hand, ۲٥۱ word families (٦۷,٤۸٪) are exclusive to the LAWL and do not appear in the AWL, emphasizing the highly specialized nature of legal terminology. This breakdown shows that while some legal vocabulary overlaps with general academic language, the majority is unique to the legal domain.

Table ٦ The initial ٥٠ words of the Law Academic Word List (LAWL) extracted from Law Research Article Corpus (LRAC) prior to the law experts' validation

Rank	Words	Frequency	Rank	Words	Frequency
١.	law	٢٨٨٥٩	٢٦.	information	٥٤٢١
٢.	court	٢٢٦٢٠	٢٧.	amendment	٤٩٦١
٣.	note	١٤٨٥٣	٢٨.	supreme	٤٨٩٩
٤.	state	١٢٨٥٥	٢٩.	process	٤٨٦٥
٥.	legal	١٠٩٨٥	٣٠.	article	٤٧٤٤
٦.	states	١٠٠٤٢	٣١.	criminal	٤٧٢٥
٧.	cases	٩٩٩٠	٣٢.	Rule	٤٦٨١
٨.	courts	٩٧٧٥	٣٣.	people	٤٤٦٣
٩.	federal	٨٧٣٧	٣٤.	police	٤٤٢٦
١٠.	data	٨٢٠٥	٣٥.	judicial	٤٤٠٢
١١.	case	٧٩٤٦	٣٦.	policy	٤٢٣٧
١٢.	public	٧٤٣٦	٣٧.	evidence	٤٣٥٢
١٣.	because	٧٥٣٧	٣٨.	civil	٤١٠٩
١٤.	rights	٧١٥٣	٣٩.	laws	٤٠٦٦
١٥.	review	٦٨٢١	٤٠.	rules	٤٠٣٧
١٦.	justice	٦٧٥٨	٤١.	judges	٤٠١١
١٧.	part	٦٤٥٤	٤٢.	decision	٤٠٨٦
١٨.	congress	٦٥١٣	٤٣.	system	٤٢٠٤
١٩.	power	٦٢٧١	٤٤.	based	٣٨٨٤
٢٠.	act	٦٢٤٨	٤٥.	executive	٣٨٣٣
٢١.	constitutional	٦١٨٧	٤٦.	political	٣٧٨٣
٢٢.	government	٥٥٩٨	٤٧.	action	٣٦٦١
٢٣.	united	٥٧٠٧	٤٨.	social	٣٦١٨
٢٤.	right	٥٦٠١	٤٩.	section	٣٥٤٤
٢٥.	**use**	٥٤٧٤	٥٠.	**authority**	٣٥٣٧

Table ٦ presents the initial ٥٠ words of the Law Academic Word List (LAWL), which were extracted from the Law Research Article Corpus (LRAC) before being reviewed and validated by law experts. This table showcases a preliminary list of words identified as potentially important in legal academic writing based on their frequency and relevance within

the corpus.These words represent a selection of terms that are commonly used in legal research articles and are likely to include both general academic vocabulary and domain-specific legal terminology. However, since this list was compiled prior to expert validation, it may still contain words that are not entirely relevant or precise for legal scholarship, or it might lack some critical terms that experts would later add.In essence, Table ٦ provides a raw, unrefined view of the language used in legal research, serving as a starting point for further refinement and validation by legal professionals to ensure the list accurately reflects the core vocabulary of the field.

Table ٧ Expert Validation Form of Law Field Terminology for Inclusion in LAWL

Word	Experts' scores					Average score	Inclusion
Law							
Court							
Note							
State							
Legal							

Table ٧ summarizes the collective evaluation of words by experts, displaying the individual scores given by each expert, the average score calculated from those ratings, and the final decision on whether each word is included or excluded based on the results. Each word was assessed by law experts, who assigned scores based on its relevance and importance in the context of legal research and writing. The experts rated the words on a scale of ١ to ٥, allowing for decimal numbers (e.g., ١,٢٥, ٢,٥٠) to provide more precise evaluations, and only those with an average score of ٣,٧٥ or higher were deemed suitable for inclusion.In summary, this table serves as a record of the validation process, showcasing how each word was scrutinized and either accepted or rejected for inclusion in the final list. It highlights the systematic approach used to ensure the selected words are highly relevant and meaningful within the field of law.

Table ٨ The top ٥٠ most common words extracted from Law Research Article Corpus (LRAC) after the Law experts' validation

Rank	Words	Frequency	Rank	Words	Frequency
١.	law	٢٨٨٥٩	٢٦.	criminal	٤٧٢٥
٢.	court	٢٢٦٢٠	٢٧.	rule	٤٦٨١
٣.	note	١٤٨٥٣	٢٨.	police	٤٤٢٦
٤.	state	١٢٨٥٥	٢٩.	judicial	٤٤٠٢
٥.	legal	١٠٩٨٥	٣٠.	evidence	٤٣٥٢
٦.	cases	١٠٠٤٢	٣١.	policy	٤٢٣٧
٧.	federal	٨٧٣٧	٣٢.	system	٤٢٠٤

۸.	data	۸۲۰۵	۳۳.	decision	٤٠٨٦
۹.	because	۷۵۳۷	۳٤.	based	۳۸۸٤
۱۰.	public	۷٤۳٦	۳۵.	executive	۳۸۳۳
۱۱.	rights	۷۱۵۳	۳٦.	social	۳٦۱۸
۱۲.	review	٦۸۲۱	۳۷.	section	۳۵٤٤
۱۳.	justice	٦۷۵۸	۳۸.	authority	۳۵۳۷
۱٤.	congress	٦۵۱۳	۳۹.	analysis	۳۵۰٦
۱۵.	part	٦٤۵٤	٤۰.	protection	۳٤۹۷
۱٦.	power	٦۲۷۱	٤۱.	question	۳۳٤٦
۱۷.	act	٦۲٤۸	٤۲.	statute	۳۳۰۸
۱۸.	constitutional	٦۱۸۷	٤۳.	claims	۳۲۸۹
۱۹.	united	۵۷۰۷	٤٤.	enforcement	۳۰۱۲
۲۰.	government	۵۵۹۸	٤۵.	private	۳۰۰۳
۲۱.	use	۵٤۷٤	٤٦.	including	۳۰۰۰
۲۲.	information	۵٤۲۱	٤۷.	litigation	۲۹۲٤
۲۳.	amendment	٤۹٦۱	٤۸.	legislative	۲۸۸٦
۲٤.	supreme	٤۸۹۹	٤۹.	doctrine	۲۸۳٤
۲۵.	**process**	٤۸٦۵	۵۰.	**approach**	۲۷۷٦

Words that coincide with Coxhead's Academic Word List are in bold font. Non-bolded words are specific to law field

Table ۸ displays the top ۵۰ most frequently occurring words from the Law Research Article Corpus (LRAC), which has been reviewed and validated by legal experts. This list represents the core vocabulary that is commonly used in academic legal writing, reflecting the key terms and concepts that are central to the field of law.The inclusion of these words was determined after a thorough analysis of the corpus, followed by expert validation to ensure that only relevant and meaningful terms were selected. This table provides insight into the language and terminology that dominates legal research, helping researchers and practitioners understand the primary focus areas and themes within legal scholarshipIn essence, this table serves as a snapshot of the most significant words in legal literature, offering a foundation for further analysis or study in the field.

Table ۹ The initial ۵۰ removed words of the Law Academic Word List (LAWL) extracted from Law Research Article Corpus (LRAC) after the law experts' validation

Rank	Words	Frequency	Rank	Words	Frequency
۱.	article	٤۷٤٤	۲٦.	practice	۲۷٦۰

۲۰.	people	٤٤٦٣	۲۷.	text	۲۷٤٦
۳۰.	civil	٤١٠٩	۲۸.	standard	۲۷۰۱
٤.	clause	۳۳۱۰	۲۹.	control	۲٦٤۲
٥.	trade	۳٤٥٦	۳۰.	fact	۲٦۳٤
٦.	like	۳٤٤۷	۳۱.	education	۲٥٦۲
۷.	history	۳٤۱۲	۳۲.	given	۲٥٥۱
۸.	make	۳٤۱۰	۳۳.	provide	۲٥٤۸
۹.	common	۳٤۰٤	۳٤.	context	۲٥٤۰
۱۰.	work	۳۲۱۸	۳٥.	market	۲٥۲۰
۱۱.	general	۳۰۹٤	۳٦.	role	۲٥۱٤
۱۲.	way	۳۰۷۷	۳۷.	meaning	۲٥۱۳
۱۳.	likely	۳۰٥۹	۳۸.	important	۲٤٦۷
۱٤.	president	۳۰٤۷	۳۹.	support	۲٤٥٦
۱٥.	including	۳۰۰۰	٤۰.	own	۲٤٥۰
۱٦.	different	۲۹۸۸	٤۱.	individuals	۲٤۳۲
۱۷.	class	۲۹۸٤	٤۲.	research	۲٤۲۰
۱۸.	Risk	۲۹۷۱	٤۳.	local	۲٤۱٦
۱۹.	Individual	۲۹٦۸	٤٤.	making	۲٤۰٥
۲۰.	Costs	۲۹٥۹	٤٥.	interest	۲٤۰۲
۲۱.	Theory	۲۸۹۹	٤٦.	test	۲۲٦٤
۲۲.	Made	۲۸۹٤	٤۷.	level	۲۲٥۱
۲۳.	international	۲۸٤٤	٤۸.	due	۲۲۲۷
۲٤.	Doctrine	۲۸۳٤	٤۹.	need	۲۲۲٥
۲٥.	**Particular**	۲۸۰۸	٥۰.	**found**	۲۲۲۲

Table ۹ presents the initial ٥۰ words that were removed from the Law Academic Word List (LAWL) after being extracted from the Law Research Article Corpus (LRAC) and subjected to validation by legal experts. Despite their presence in the corpus, all the words listed here were excluded from the final word list because they received an average expert score of less than ۳،۷٥ out of ٥ , failing to meet the threshold for relevance and importance in legal scholarship.This table highlights the rigorous selection process used to refine the LAWL, demonstrating how even frequently occurring words can be excluded if they do not meet the high standards set by legal experts. It emphasizes the importance of balancing word

frequency with contextual significance in creating a specialized vocabulary list for legal education and research.

The Law experts' opinions toward Law Academic Word List

However, after the researcher complied this LAWL, the validation form with this word list was sent to five law experts for validation. The experts were asked to go through the words on the list and decide whether or not each word should appear on the list. A word would be removed if its mean score was lower than ۰٫۷۵. The experts employed a rating scale spanning from ۱ to ۵, encompassing the categories of Not valid, Slightly valid, Moderately valid, Very valid, and Highly valid. A numerical value would be assigned to each score (e.g. ۱=۱, ۲=۲, ۳=۳, ٤=٤, ۵=۵). Then, for each word, we would add up the scores given by all five experts and divide by ۵ to get the mean score. For example, if expert ۱ gave a word a score of ٤, expert ۲ gave it a score of ۵, expert ۳ gave it a score of ۳, expert ٤ gave it a score of ٤, and expert ۵ gave it a score of ۲, we would add up ٤+۵+۳+٤+۲=۱۸ and divide by ۵ to get a mean score of ۳٫٦. In this scoring system, the highest mean score is ۵.The words with a mean score less than ۰٫۷۵(۳٫۷۵ of ۵) will be removed.Experts can use decimal scores (e.g., ۱٫۲۵, ۲٫۵۰, ۳٫۷۵, ٤٫۲۵) to provide precise and nuanced evaluations of each word's validity. This approach captures subtle distinctions in their judgments, avoiding the limitations of whole numbers. For example, a score of ۳٫۷۵ indicates a word is closer to "very valid" (٤) than "moderately valid" (۳), while ۲٫۵۰ places it midway between "slightly valid" (۲) and "moderately valid" (۳). Decimal scores enhance the validation process by making it more precise, fair, and reflective of the experts' true opinions, ultimately yielding a more accurate and reliable final word list.

Discussion

The study aimed to accomplish two primary objectives: firstly, to construct a Law Academic Word List (LAWL), and secondly, to investigate the degree of overlap between the words in the LAWL and those included in Coxhead's academic word list. The approach involved the compilation of the Law Research Article Corpus (LRAC), a collection of ۱۲۰۰ articles sourced from the top ۱۰ law journals, containing a total of ٤,۱۱٤,۳۲۸ running words. Subsequently, the Software AntWordProfiler was employed to carry out data analysis and determine word frequency, while excluding certain non-relevant elements such as graphs, charts, bibliographies, text headers, footnotes, author names, generic names, family names, and function words.

The initial count of words in the Law Academic Word List, Stood at ۳۷۳۰ words.The validation process of the LAWL was conducted by five experts in the field of law, culminating in the development of the LAWL. ۲۵٤۱words were removed during the validation process. The removed words were those for which the validation criteria were not met or which were deemed unsuitable for inclusion in the final list. Number of ۱۱۸۹ words passed the validation process successfully. These words met the necessary criteria and were deemed valid for inclusion in the finalized academic word list. The study's findings revealed that Coxhead's Academic Word List only covered ۳۲٫۵۲٪ of the terms present in the LAWL, signifying that a considerable ٦۷٫٤۸٪ of the words in the LAWL were distinct to the field of law and were not

encompassed by Coxhead's list. This disparity implies the specialized nature of the vocabulary within the field of law, suggesting the existence of a substantial body of legal terminology that is not adequately represented in more generalized academic word lists. The development and validation of the LAWL hold particular significance for academic and professional contexts within the legal domain, aiding in the comprehension, communication, and study of legal discourse. The study's results underscore the necessity of domain-specific linguistic resources and the importance of recognizing the unique lexical demands of specialized fields such as law. Further research might be necessary to examine the implications of these findings for language learning, legal scholarship, and communication within the legal profession.These findings offered implications for language teachers and material designers to establish discipline-specific word lists. This will be beneficial to English language learners in order to acquire the words effectively and for their own field of study. Teachers and course designers should have reference word lists to decide which words in a certain register should be taught to students (Nation and Waring, ۱۹۹۷). Particularly, the Law Academic Word Lists aimed to be a ready-to-use word list for instructors and course designers of English for Academic Purposes (EAP) and English for Specific Purposes (ESP) in law. Overall, this endeavor is another application of using corpus linguistics and the empirical evidence that will enhance students' learning.

References

Adolphs, S. & Lin, P. M. S. (۲۰۱۱). Corpus Linguistic. In J. Simpson (Ed.), The Routledge Handbook of Applied Linguistics (pp. ۵۹۷-۶۱۰). Routledge.

Afsharzadeh, A. (۲۰۲۱). Developing a corpus-based academic word list and collocation list in medicine. Journal of Language and Communication, ۸(۲), ۲۳۹-۳۰۰.

Anthony, L. (۲۰۱٤b). AntWordProfiler (Version ۱،٤. ۱)[computer software]. Tokyo, Japan

Baumann, J. F., & Graves, B. (۲۰۱۰). Teaching reading in today's world: A balanced approach. Pearson

Bawarshi, Anis S. & Mary Jo, Reiff. (۲۰۱۰). *Genre: An Introduction to History, Theory, Research, and Pedagogy.* Parlor Press; The WAC Clearinghouse. https://wac.colostate.edu/books/referenceguides/bawarshi-reiff/

Becher, T., & Trowler, P. (۲۰۰۱). Academic tribes and territories. intellectual enquiry and the culture of disciplines (۲nd ed.). Buckingham, UK: The Society for Research into Higher Education & Open University Press.

Bhatia, V. K. (۲۰۰۲). Applied Genre Analysis: Analytical Advances and Pedagogical Procedures. In A. M. Johns (Ed.), Genre in the Classroom: Multiple Perspectives (pp. ۲۷۹-۲۸۳). Mahwah, NJ: Erlbaum.

Bhatia, V. K. (۲۰۱۷). Critical genre analysis: Investigating interdiscursive performance in professional practice. London, UK: Routledge.

Biber, D., Conrad, S., & Reppen, R. (۱۹۹۸). Corpus Linguistics: Investigating Language Structure and Use. New York: Cambridge University Press. http://dx.doi.org/۱۰،۱۰۱۷/CBO۹۷۸۰۵۱۱۸۰٤٤۸۹

Brezina, V., & Gablasova, D. (۲۰۱۵). A new general service list: A corpus-based approach to the identification of high-frequency words in English. International Journal of Corpus Linguistics, ۲۰(۱), ۱-۲۲.

Browne, C., Culligan, B., & Phillips, J. (۲۰۱۳). The new general service list: A core vocabulary for EFL students and teachers. JALTs The Language Teacher, ۳٤(۷), ۱۳-۱۵.

Bruce, I. (۲۰۰۸). Academic writing and genre: A systematic analysis. New York, NY: Continuum International Publishing Group.

Campion, M., & Elley, W. (۱۹۷۱). An academic vocabulary list. Wellington: New Zealand Council for Educational Research.

Carter, D. (۱۹۸۳). Some Propositions about ESP. The ESP Journal, ۲, ۱۳۱-۱۳۷.

Charles, M. and Pecorary, D. (۲۰۱٦). Introducing English for Academic Purposes. NY. Routledge.

Chen, Y., & Ge, N. (۲۰۰۷). The Academic Word List in medical research articles. Journal of English for Academic Purposes, ٦(۳), ۱۸۳-۱۹٤.

Cheng, W. (۲۰۱۲). Exploring Corpus Linguistics: Language in Action. Routledge.

Coady, J., Magoto, J., Hubbard, P., Graney, J., & Mokhtari, K. (۱۹۹۳). High Frequency Vocabulary and Reading Proficiency in ESL Readers. In T. Huckin, M. Haynes, & J. Coady (Eds.), Second Language Reading and Vocabulary Learning (pp. ۲۱۷-۲۲۸). Norwood, NJ: Ablex Publishing Corporation.

Cobb, T., & Horst, M. (۲۰۰٤). Is there room for an academic word list in French? Vocabulary in second language, ۲۳, ۱٥-۳۸.

Corson, D. (۱۹۹۷). The learning and use of academic English words. *Language Learning*, *47*(٤), ٦۷۱–۷۱۸. https://doi.org/۱۰,۱۱۱۱/۰۰۲۳-۸۳۳۳,۰۰۰۲٥

Coxhead, A. (۲۰۰۰). A new academic word list. TESOL Quarterly, ۳٤(۲), ۲۱۳–۲۳۸.

Coxhead, A. (۲۰۱۱). The Academic Word List ۱۰ Years On: Research and Teaching Implications. TESOL Quarterly, ٤٥ (۲), ۳٥٥–۳٦۲. https://doi.org/۱۰,٥٥٤/tq.۲۰۱۱,۲٥٤٥۲۸

Coxhead, A., & Hirsch, D. (۲۰۰۷). A pilot science-specific word list. Revue française de linguistique appliquée, ۱۲(۲), ٦٥-۷۸.

Coxhead, A., & Nation, P. (۲۰۰۱). The specialized vocabulary of English for academic purposes.

Farrell, T. J. (۱۹۹۰). Norms of rhetorical practice in the sciences. Written Communication, ۷(۱), ۱–۲۳.

Flowerdew, L. ۲۰۱۲. Corpora and Language Education. Basingstoke: Palgrave Macmillan. (xv + ۳٤۷pp.)

Francis, N. W. (۱۹۹۲). Language Corpora B. C.. In J. Svartvik (ed.), *Directions in Corpus Linguistics. Proceedings of the Nobel Symposium 82* , Vol. ٦٥ (pp. ۱۷–۳۲) . Mouton de Gruyter .

Gardner, D., & Davies, M. (۲۰۱۳). A new academic vocabulary list: Based on the corpus of contemporary American English. Routledge.

Gardner, D., & Davies, M. (۲۰۱٤). A new academic vocabulary list. Applied linguistics, ۳٥(۳), ۳۰٥-۳۲۷.

Gee, J. P. (۱۹۹٦). Social Linguistics and Literacies: Ideology in Discourse. London: Taylor and Francis.

Ghadessy, P. (۱۹۷۹). Frequency counts, word lists, and materials preparation: A new approach. English Teaching Forum, ۱۷, ۲٤–۲۷.

Gilner, L., & Morales, F. (۲۰۰۸). Elicitation and Application of a Phonetic Description of the General Service List. System, ۳٦(٤), ٥۱۷-٥۳۳.

Hartley, J. (۲۰۰۸). Academic writing and publishing: A practical handbook. New York: Taylor & Francis.

Heidari Tabrizi, H., & Jamalzadeh, M. (۲۰۲۰). Academic Vocabulary in Tourism Research Articles: A Corpus-Based Study. Journal of Language and Discourse Practice, ۱(۲), ۲٥-٤٤.

Hewings, M. (۲۰۰۱). Academic writing in context: Implications and applications: Papers in honor of Tony Dudley-Evans. Birmingham: University of Birmingham Press.

Hiebert, E. H., & Lubliner, S. (۲۰۰۸). Teaching reading comprehension: A resource book. Guilford Press.

Hlavacova, J. (۲۰۰٦). Average Reduced Frequency: A new measure for lexical analysis. Journal of Language and Linguistic Studies, ۲(۱), ٤٥-٦۰.

Hsu, W. (۲۰۱۱). A business word list for perspective EFL business postgraduate. The Asian ESP Journal, ۷(٤), ٦۳-۹۹.

Hutchinson, T., & Waters, A. (۱۹۸۷). English for Specific Purposes: A Learner-Centered Approach. Cambridge: Cambridge University Press.

Hyland, K. (۲۰۰٦). English for academic purposes: An advanced resource book. London: Taylor & Francis.

Hyland, K. (۲۰۰۹). Corpus informed discourse analysis: The case of academic engagement. In M. Charles, D. Pecorari, & S. Hunstan (Eds.), Academic writing: At the interface of corpus and discourse (pp. ۱۱۰–۱۲۸). New York: Continuum International Publishing Group.

Hyland, K. (۲۰۱۲). Bundles in academic discourse. Annual Review of Applied Linguistics, ۳۲, ۱۵۰-۱۶۹. https://doi.org/۱۰,۱۰۱۷/S۰۲۶۷۱۹۰۵۱۲۰۰۰۰۳۷

Hyland, K., & Tse, P. (۲۰۰۷). Is there an 'Academic vocabulary'? TESOL Quarterly, ٤۱(۲), ۲۳۵–۲۵۳.

Hyland, K., Huat, C. M. and Handford, M. (۲۰۱۲). Corpus Applications in Applied Linguistics. London: Bloomsbury

Jalilifar, A. (۲۰۱٤). Directions in discourse analysis: Theory and method. Ahvaz: Shahid Chamran University Press.

Juilland, A., Brodin, P., & Davidovitch, A. (۱۹۷۰). A quantitative study of French vocabulary and verbal productivity. Mouton.

Khani, R., & Kalantari, B. (۲۰۲۱). Accounting Academic Word List (AAWL): A Corpus-Based Study. Journal of Foreign Language Teaching and Translation Studies, Vol.٦, No.۱, Winter ۲۰۲۱, pp.۳۵-۵۸.

Khani, R., & Tazik, K. (۲۰۱۳). Towards the development of an academic word list for applied linguistics research articles. RELC journal, ٤٤(۲), ۲۰۹-۲۳۲.

Konstantakis, N. (۲۰۰۷). Creating a business word list for teaching business English. ELIA, ۷, ۷۹-۱۰۲.

Konstantakis, N. (۲۰۰۷). The usefulness of the academic word list for learners of English for academic purposes in higher education. Journal of English for Academic Purposes, ٦(٤), ۲۸۸-۳۰۲.

Konstantakis, N. (۲۰۱۰). Academic word list: A valuable resource for language learners. International Journal of Lexicography, ۲۳(۳), ۲۳۹-۲۵۷.

Lam, J. K. M. (۲۰۰۱). A study of semi-technical vocabulary in computer science texts, with special reference to ESP teaching and lexicography.

Lee, D. Y. W. (۲۰۰۸). Corpora and discourse analysis: New ways of doing old things. In V. K. Bhatia, J. Flowerdew, & R. H. Jones (Eds.), Advances in discourse studies (pp. ۸٦–۹۹). London: Routledge.

Lei, J., & Liu, D. (۲۰۱٦). A critical review of Gardner and Davies's (۲۰۱٤) new academic vocabulary list. Journal of English for Academic Purposes, ۲۳, ۱-۱۰.

Li, D., & Qian, Y. (۲۰۱۰). The coverage of academic vocabulary in textbooks for English for academic purposes. Journal of English for Academic Purposes, ۹(۳), ۱۸۵-۱۹٤.

Li, E. & Pemberton, R. ۱۹۹٤. An investigation of students' knowledge of academic and subtechnical vocabulary. In Joint seminar on corpus linguistics and lexicography(p. ۱۸۳-۱۹٦). Hong Kong: HKUST Language Center.

Liu, J., & Han, L. (۲۰۱۵). A corpus-based environmental academic word list building and its validity test. English for Specific Purposes, ۳۹, ۱–۱۱.

Lynn, R. W. (۱۹۷۳). Preparing word lists: a suggested method. RELC Journal, ٤(۱), ۲۵–۳۲.

Martínez, I. A., Beck, S. C., & Panza, C. B. (۲۰۰۹). Academic vocabulary in agriculture research articles: A corpus-based study. English for specific purposes, ۲۸(۳), ۱۸۳-۱۹۸.

McCarthy, M. (۲۰۱۹). Why do we need EAP word lists and what can we do with them. In BALEAP Conference: Innovation, Exploration and Transformation, University of Leeds ۱۲th–۱٤th April ۲۰۱۹.

McEnery, A. M., & Wilson, A. (۱۹۹٦). Corpus Linguistics (Edinburgh Textbooks in Empirical Linguistics Series) (p. ۲۰۹). Edinburgh University Press.

McKinley, J., & Rose, H. (۲۰۱۸). Conceptualizations of language errors, standards, norms and nativeness in English for research publication purposes: An analysis of journal submission guidelines. Journal of Second Language Writing ٤۲, ۱-۱۱.

Miller, G. A., & Biber, D. (۲۰۱۵). Lexical diversity in academic writing: A comparison of the New General Service List and the General Service List. Journal of English for Academic Purposes, ۲۰, ۵۰-۵۹. https://doi.org/۱۰,۱۰۱٦/j.jeap.۲۰۱۵,۰۷,۰۰۱

Montgomery, S. (۱۹۹٦). The scientific voice. New York: The Guildford Press.

Mudraya, O. (۲۰۰۶). A corpus-based study of academic vocabulary in engineering textbooks. Journal of English for Academic Purposes, ۵(٤), ۲۵۹-۲۷۲.

Nagy, W., & Townsend, D. (۲۰۱۲). Words as tools: Learning academic vocabulary as language acquisition. *Reading Research Quarterly, 47*(۱), ۹۱–۱۰۸. https://doi.org/۱۰,۱۰۰۲/RRQ.۰۱۱

Nation, I. S. P. (۲۰۰۱). Learning vocabulary in another language. Cambridge University Press.

Nation, I. S. P. (۲۰۰٤). A study of the General Service List. Language Testing, ۲۱(٤), ۳۹۳-٤۱۸. https://doi.org/۱۰,۱۱۹۱/۰۲٦٥٥۳۲۲۰٤lt۲۸۷oa

Nation, I. S. P. (۲۰۱٦). Making and using word lists for language learning and testing. Making and Using Word Lists for Language Learning and Testing, ۱-۲۲٤.

Nation, P., & Hwang, K. (۱۹۹٥). Where would general service vocabulary stop and special purpose vocabulary begin? *System, 23*, ۳٥-٤۱. [-۲۳-]

O'Flynn, J. A. (۲۰۱۹). An Economics Academic Word List (EAWL): Using online resources to develop a subject-specific word list and associated teaching-learning materials. Journal of academic language and learning, ۱۳(۱), A۲۸-A۸۷.

O'Keeffe, A. & McCarthy, M.(۲۰۱۰) 'Historical perspective: What are corpora and how have they evolved?', in O'Keeffe, A. & McCarthy, M.(eds), The Routledge Handbook of Corpus Linguistics. London: Routledge, p ۳-۱۳.

O'Keeffe, A., McCarthy, M. J. and R. A. Carter (۲۰۰۷) From Corpus to Classroom: language use and language teaching. Cambridge: Cambridge University Press.

Paquot, M. (۲۰۱۰). Academic vocabulary in learner writing: From extraction to analysis. Bloomsbury Publishing.

Praninskas, J. (۱۹۷۲). American university word list. London: Longman.

Read, J. (۱۹۹۸). Validating a Test to Measure Depth of Vocabulary Knowledge. In A. Kunnan (Ed.), Validation in Language Assessment (pp. ٤۱-٦۰). Mahwah, NJ: Erlbaum.

Renouf, A. (۱۹۸۷). Corpus development Looking up: An account of the COBUILD Project in lexical computing, ۱-٤۰.

Richards, J. C. (۱۹۷٤). A Non-Contrastive Approach to Error Analysis. In Error Analysis: Perspectives on Second Language Acquisition (pp.۱۷۲-۱۸۸). London: Longman.

Rorty, R. (۱۹۷۹). Philosophy and the mirror of nature. Princeton: Princeton University Press.

Savicky, P., & Hlavacova, J. (۲۰۰۲). The role of frequency in language learning: A study of word usage in corpora. *Language Learning*, ٥۲(٤), ۸۹۳-۹۱٤. https://doi.org/۱۰,۱۱۱۱/۱٤٦۷-۹۹۲۲,۰۰۲۰۲

Shirazi, D., & Dolatabadi, H. R. (۲۰۲۰). Developing an Academic Word list for the students of Health Information Management: A Corpus study. Iranian Journal of English for Academic Purposes.

Swales, J. (۱۹۹۰). Genre analysis: English in academic and research settings. New York: Cambridge University Press.

Swales, J.M. (۲۰۰٤) Research Genres: Exploration and Application. Cambridge University Press, Cambridge. http://dx.doi.org/۱۰,۱۰۱۷/CBO۹۷۸۱۱۳۹٥۲٤۸۲۷

Valipouri, L., & Nassaji, H. (۲۰۱۳). A corpus-based study of academic vocabulary in chemistry research articles. Journal of English for Academic Purposes, ۱۲(٤), ۲٤۸-۲٦۳.

Vongpumivitch, V., Huang, J. Y., & Chang, Y. C. (۲۰۰۹). Frequency analysis of the words in the Academic Word List (AWL) and non-AWL content words in applied linguistics research papers. English for Specific Purposes, ۲۸(۱), ۳۳-٤۱. https://doi.org/۱۰,۱۰۱٦/j.esp.۲۰۰۸,۰۸,۰۰۳