Applied Linguistics and Urban Planning Nexus: Developing an Academic Word List of Urban Planning Using Corpus Linguistics Approach

Mohammad Amini Farsani ✉* ⁣📙 0000-0002-0249-1996
Department of Foreign Languages, University of Science and Technology, Tehran, Iran.
Email: Mohammad_Farsani@iust.ac.ir

Mohammad Hossein Afshari ** 📙 0009-0000-6989-6712
Department of Foreign Languages, University of Science and Technology, Tehran, Iran.
Email: **Mehran.Afshari.Tefl@gmail.com**

Taha Saedi Roudi* 📙 0009-0004-7557-5965
Department of Foreign Languages, University of Science and Technology, Tehran, Iran.
Email: Taha.Saedi.Acad@gmail.com

ABSTRACT

Having authentic, technical, and field-specific word lists at the disposal of each academic community is crucial and yet contemporary for the various academic obstacles that need to be addressed in specific fields. Given the numerous vocabularies within a language, especially specific language of a field, it is vital to provide students with vocabularies that they are most likely to encounter with in academic discourses. As such, this study was aimed at establishing a specialized corpus to identify the most prevalent academic words, leading to the development of Urban Planning Academic Word List (hereafter UPAWL). To do so, researchers identified top-tier journals of Urban Planning and included four journals with the highest elite scores (multiplying Impact Factor by Citation Score). To include the research articles of Urban Planning, research articles were required to follow the IMRD format (Swales, 1990) and be published from 2013-2017 and 2018-2022. By applying simple random sampling as well as stratified random sampling, we included 100 articles, bringing about an Urban Planning corpus with 1,252,458 million running words. To develop the academic word list, we followed the Coxhead (2000) criteria, including range, frequency, and specialized occurrence. This process led to identification of 507 word types. The implications for researchers, students and EAP instructors of Urban Planning are also discussed.

✉* TEFL, Department of Foreign Languages, University jf Science and Technology, Tehran, Iran.
** TEFL, Department of Foreign Languages, Iran University of Science and Technology, Tehran, Iran.
*** TEFL, Department of Foreign Languages, Iran University of Science and Technology, Tehran, Iran.

697

Introduction

With the rise of globalization and world Englishes, English has increasingly become the dominant language for communication in academic discourses (Alhasnawi, 2021). This global function of English is most likely to open the path to a multitude of scholarly informational resources, such as scientific research, historical documents, English for Academic Purposes (EAP) and English for Specific Purposes (ESP) research products. As a result, this linguistic paradigm allows students, researchers, instructors, and scholars to engage with cutting-edge research, evidence-based findings, research-informed materials, and textbooks far beyond their native language. More importantly, within this paradigm, the globalization of English serves as a bridge for international collaboration, in which different people with various backgrounds can communicate smoothly and exchange ideas, data, information, and findings dynamically, since it facilitates effective communication among representatives of different countries, continents, and regions (Kopchark et al., 2022). For example, in EAP and ESP strands of applied linguistics, the affordances of English in today's integrated world are not only a scholastic tool, but also a prerequisite for entering the territory of the academic community.

One of the often-cited requirements to enter the arena of academia is to represent the language and nomenclature of general and specific disciplines through developing and establishing academic word lists

(Gholaminejad & Anani Sarab, 2020). This requirement has been empirically manifested in studies focused on developing general word lists, such as GSL[1](West, 1953), vocabularies commonly used in everyday conversations, reading materials, newspapers, magazines, official documents, and writing tasks. Likewise, Coxhead (2000) developed a word list of academic words, named AWL[2],which was extracted from research articles of four scientific areas: commerce, law, arts, and science. This could support the point that was made by Hyland and Tse (2007) where they believed that the unequal distribution of academic words in each discipline heightens the need to develop discipline-specific word lists. This word list, targeting hard and soft sciences, provided evidence-based academic words for a variety of disciplines and majors such as accounting (Khany & Kalantari, 2021), chemistry (Valipouri & Nassaji, 2013; Xodabandeh et al., 2023), computer science (Roseler, 2021) physics (Milica Vukovic-Stamatovic,2024), veterinary (Özer & Akbas, 2024), linguistics (Moini & Islamizadeh, 2016), applied linguistics (Barraza et al., 2024; Farahani et al., 2020; Matinparsa et al, 2023), psychology(Xodabandeh et al., 2020), aerospace(Korzin et al., 2023), especially designed to address linguistic needs of non-native English speakers.

The purpose of developing these academic word lists was to enable users to meet their linguistic needs in daily life and academic discourses, respectively. However, Coxhead (2000) acknowledged that AWL cannot serve

---

students' needs who major in different areas of science equally, since the extent of the provided coverage of AWL was not equal in different areas of science. As Hyland and Tse (2007) put it, different lexical items demonstrated different lexical behaviors across a diversity of scientific areas in terms of meaning, range, and collocation. This phenomenon highlights the significance of developing field-specific word lists, tailored for specific scientific areas. In this study, we aim to develop discipline-specific word lists for the field of urban planning. This study was guided by the following research questions:

I) What are the most frequent academic words of Urban Planning?

II) To what extent do GSL and AWL cover the entire corpus of Urban Planning Research Articles?

Review of Literature

Nation (2001) identified four distinct vocabulary categories prevalent in English academic writing: high-frequency words, academic words, technical words, and low-frequency words. According to such vocabulary categorization, based on general corpora of BNC[1]/COCA[2], vocabularies that were in the top 3 levels (each consisting of 1000 words) of BNC/COCA are high-frequent. The fourth level to the nineth level of BNC/COCA form the mid-frequent vocabularies and vocabularies that go beyond the 9 levels of BNC/COCA form the low-frequent words (Nation, 2012). High-frequency words are commonly used in everyday conversations, reading materials,

and writing tasks. Numerous word lists have been developed to facilitate students' acquisition of essential and significant vocabulary, with West's (1953) General Service List of English Words (GSL) being the most widely utilized and recognized compilation of high-frequency words. On the other hand, low-frequency words, characterized by their infrequent usage and limited distribution, constitute a considerable portion of any given field's vocabulary. Some of these words may appear only once or twice, encountered sporadically. Nevertheless, they form the largest group of words within the field. Low-frequency words account for approximately 5% of the vocabulary found in academic texts, encompassing proper names, words seldom encountered in general language usage, non-high-frequency words, and technical vocabulary from other subject areas (Nation, 2001). As aptly stated by Nation (2001), "one person's technical vocabulary is another person's low-frequency words," underscoring the variability which is inherent in low-frequency vocabulary.

Academic words are not typically encountered in basic general English texts but instead constitute a relatively substantial proportion of the vocabulary found in academic discourse. Learners often face more challenges in acquiring these words as they are less familiar with them compared to the technical vocabulary specific to their respective fields. The Academic Word List (Coxhead, 2000), frequently referenced in literature, is comprised of 570-word families

---

[1] British National Corpus

[2] Corpus of Contemporary American English

that fall outside the realm of the 2,000 most frequently used English words. It serves as a valuable learning resource for learners with academic objectives (Coxhead & Nation, 2001). As Liu and Han pointed out (2015, p.1), academic word lists can be classified into two types: general academic word lists and field-specific academic word lists. The former refers to word lists which are composed of words related to various disciplines and can be acquired and used by the majority of ESP students as a prerequisite for their university studies, such as AWL (Coxhead, 2000), while the latter insinuates to words commonly found in different subject areas of a particular discipline (Martinez et al, 2009; Khani & Tazik, 2013). In contrast, technical words, pertain to specialized topics, fields, or disciplines. While they may be relatively common within a particular subject area, they are encountered less frequently in other contexts. Technical words encompass a diverse range of types, some of which are unique to specific subject domains (Nation, 2001).

Previous Studies

Given the paramount significance of technical word lists and AWL's inadequacy to meet linguistic needs of EAP students across diversity of fields equally, several studies have been conducted to identify the academic vocabularies of scientific disciplines. As such, Wang et al. (2008) aimed to develop a word list for medical students to investigating lexical profile of technical vocabularies of medical science, adopting a corpus-based approach. As the initial stages of their study, they used ScienceDirect online

(http://www.sciencedirect.com), as their database to choose both population of their study, academic journals of Medical Science, and areas in which they were going to cover. This database subsumes almost every top title across twenty-four disciplines, from natural sciences to social sciences. Following the previous stage, they identified thirty-two areas within the discipline of medicine and dentistry. Considering each of the identified areas as one stratum, researchers singled out three journals from each area using stratified random sampling approach. Ultimately, ninety-six journals were extracted out and one issue, out of the chosen journals, was selected randomly. Identifying the population of the study, researchers defined some inclusion criteria in that research articles had to follow the IMRD (Introduction, Methods, Results, and Discussion) format (Swales,1990), the strict rule (i.e., The first author of the selected articles should be native English speakers or should have affiliations with the organizations in native English speaking countries) (Wood, 2006), word length framework of 2000 to 12000, and publishing framework from 2000 to 2007. Considering the criteria, Wang et al. (2008) selected 288 research articles and established a corpus with 1.093.011 running words; by doing so, 31257-word families were identified, and medical science academic word list was developed. Applying word selection criteria, proposed by Coxhead (2000), specialized occurrence, range, and word frequency, researchers developed a list of specialized academic vocabulary for medical writing. As for the word frequency, the chosen words had to appear at least 30

times in the entire corpus, be used in at least 16 of the 32 subject areas for range and fall outside of the top 2000 most frequently used English words of GSL (West, 1953) to meet the criterion of specialized occurrence. The study identified 623-word families that met the aforementioned criteria, providing a coverage around 12.24% in the entire corpus.

In another study, Hsu (2013) carried out a study in order to bridge the gap between non-technical and technical words of medical science. After identification of subject areas of medicine and dentistry, using ScienceDirect online, the researcher chose five medical textbooks based on the identified areas. Therefore, 155 medical textbooks were chosen across 155 medical subject areas from online resources. By doing so, Hsu managed to establish a corpus with 15,016,553 running words. As for the word selection criteria, Hsu followed the defined criteria of Coxhead (2000), frequency, range, and specialized occurrences. As for the specialized occurrences, the included word families had to be outside the top BNC 3,000 most frequently occurring words of English. Given the range of word families, their member had to occur at least in more than a half of the thirty-one medical subject areas and as for the frequencies of the words, members of a word family, taken together, had to occur at least 863 times in the Medical Textbook Corpus. Applying the inclusion criteria upon identified word families, the researcher chose 595 core medical academic words in which cover about 10.72% of the whole corpus. The primary differentiation between the study conducted by Hsu (2013) and that of Wang et al. (2008) lies in their

701

respective focal points; the former investigation primarily centers on medical textbooks, whereas the latter concentrate on research articles to construct a field-specific word list for medical science. Notably, Hsu's examination identified 595 core medical academic words, with only seventy-six of these words demonstrating synchronization with Coxhead's Academic Word List (AWL). Conversely, Wang et al. (2008) identified 623 core medical academic words, with 342 of these overlapping with Coxhead's AWL. This discrepancy underscores the distinct approaches and outcomes of the two studies regarding the integration of academic vocabulary within the medical domain.

Kongnui and Phoocharoensil (2021) developed a field-specific word list for musicology. As their first stage, they selected nine Q1 top tier journals of musicology and one expert recommended journal. Three hundred research articles were derived from the selected journals which were published during 2015-2021. All the research articles were in word length range of 5300 to 8000. Compiling and analyzing three hundred research articles of musicology, researchers developed a specialized corpus of musicology with two million running words. As for the words to be included within the musicology word lists, the researchers adapted the specified outline of Coxhead (2000), frequency and specialized occurrence. As for the specialized occurrence, the words had to be outside of GSL and AWL and to meet the criterion of frequency, the words had to have fifty-seven times occurrence across the whole corpus. Given the inclusion criteria of musicology

word list (hereafter MWL), 516-word families were identified in which covered 16.44% of the entire corpus. Meanwhile, GSL (West, 1953) and AWL (Coxhead, 2000) provided a coverage of 73.25% and 10.31% in the whole corpus, respectively. As the last stage of developing the MWL, the researchers consulted three musicologists to validate the identified frequent words of musicology. As a result, 378 of the identified word families were removed from the list and the final version of MWL was developed.

Mudraya (2006) intended to develop a field-specific word list for engineering students to meet EAP and ESP objectives of engineering students of Walailak University. As the first stage of establishing an engineering corpus, she identified thirteen engineering textbooks which cover engineering areas which are compulsory for all engineering students to study, regardless of their special area of science which they study. The included areas of engineering within her study were Engineering Mechanics, Engineering Materials, Mechanics of Materials, Mechanics of Fluids, Thermodynamics, Electrical Engineering, Engineering Drawing, Manufacturing Process and Computer Programming. Compiling and analyzing thirteen textbooks of engineering, the Student Engineering English Corpus (hereafter SEEC) with nearly two million running words was constructed. Creating the SEEC, she identified 1200-word families and 9000-word types, frequently used within the corpus. In order for word families to be qualified for inclusion within SEEC word list, they should have the occurrence of sum

total of one hundred times or 0.005%. Thus, 1260-word families were included within the word list of SEEC. As the subsequent stage of developing SEEC, she performed a word frequency analysis, being concerned with the most frequent words in different corpora, to compare SEEC to COBUILD, Bank of English Corpus and BNC. The correlation between the fifty most frequent closed-class word forms in the SEEC, the COBUILD, Bank of English and the BNC Written proved to be statistically significant at the .01 level. The Spearman's rank order correlation between the fifty most frequent closed-class word forms in the SEEC and the COBUILD Bank of English is .778 while between the SEEC and the BNC Written it is .802.

Not being the only one who was interested in analyzing linguistics features of Engineering and developing a field-specific word list, Ward (2009) conducted a follow up study of Mudraya's, aiming to develop an engineering word list in order to meet students' EAP needs. In order to do so, he consulted lecturers in five engineering faculties, chemical engineering, civil engineering, electrical engineering, industrial engineering and mechanical engineering. Each of the lecturers were asked to introduce names of five commonly used textbooks in third and fourth year of undergraduate students. Therefore, twenty-five textbooks were compiled and then, some random pages were singled out till they reached 10000. By doing so, he utilized a sizable corpus of 271,000 words, with 10.290 identified word types, to construct a foundational word list tailored to the needs of engineering students. The Basic Engineering List (BEL),

consisting of 229 words, aimed to facilitate English language learning for low-level learners across all engineering disciplines. Ward astutely criticized the shortcomings of high school education in equipping engineering students with the necessary English language skills to navigate college-level textbooks, leading to conduction of an evaluation of the Basic Engineering Word List (hereafter BEL) against three other prominent undergraduate textbooks within the respective field. The assessment revealed that BEL encompassed 17.2%, 15.6%, and 21% of the sub corpora. Moreover, when compared to a text focusing on mass transfer, which pertains to both chemical and mechanical engineering, BEL exhibited a coverage of 17.7%. This observation suggests consistent and substantial coverage provided by BEL across diverse technical materials. Furthermore, while the Academic Word List (AWL) comprises a significantly larger vocabulary compared to BEL, it yielded only 11.3% coverage of engineering content. This juxtaposition highlights BEL's efficacy in providing extensive coverage of engineering terminology, despite its narrower focus compared to broader academic word lists. It is worth noting that although BEL lacks technical breadth, its distinct characteristic lies in its emphasis on engineering-specific terminology.

While the conducted studies upon the linguistic features of the field of engineering made contributions to the literature, none considered the significance of the academic genre of research articles and were mainly focused on textbooks. Also, one of them carried out more of a localized study, aiming

703

to tailor linguistic needs of EAP students in Walailak University while the other corpus was not quite reliable due to the small size of the corpus. This study is unique in the sense that it focuses exclusively on one major within engineering—Urban Planning, thereby yielding more robust and reliable results. It bears similarities to the studies conducted by Mudraya (2006) and Ward (2009), which also aimed to investigate the lexical profiles of engineering fields. However, unlike these studies, which included textbooks, the present research specifically analyzes research articles as the academic genre of focus.

Methods

Identifying the scientific area of Urban Planning, we followed the proposed benchmark of Plonsky (2013,2014), to operationalize the domain. It includes content (scope of research articles), location (i.e., journals of Urban Planning and research articles), and time (date of publication). To establish the specialized corpus of urban planning, first we aimed to identify Q1 journals of the field, using SJR and employing multilayered sampling approach and defining the elite score, multiplying their impact factor by their citation score (AUTHOR, 2021, p.4). Then, four of the journals were randomly chosen and validated by experts (See Table 1). Identifying top journals of the field, we included empirical research articles that followed the IMRD format (Swales, 1990) and were published from 2013-2017 and 2018-2022, with the purpose of achieving a representative corpus across different years. Also, to establish a balance within the corpus, the same number

of research articles were included to establish the corpus. Therefore, a specialized corpus of urban planning was constructed with 1,252,458 running words. Previous research suggests that a corpus of one million running words is sufficient to obtain a reliable list of highly frequent words (Brysbaert & New, 2009).

Table 1

Journals' Information

Journal	IF	CS	ES	Time-span
Urban Studies	4.2	10.5	44.1	2018-2022
Journal of Urban Economics	5.7	10.6	60.4	2018-2022
Progress in Planning	5	10.7	53.5	2013-2017
Environmental and Planning A	4.6	9.5	43.7	2013-2017

Corpus Establishment

To construct the corpus, first, all the research articles went under the process of standardization where all the references, footnotes, tables, figures, and appendices were removed; then, they were all converted into txt files, each txt file representing one journal, to be inserted into the concordancer of AntWordProfiler 1.5.1(Anthony, 2018).

Word Selection Criteria

To develop the field-specific word list of urban planning, we followed the proposed criteria of Coxhead (2000), specialized occurrence, range, and frequency. To identify the technical vocabularies of the field, they had to go beyond the General Service List and Academic Word List, securing and assigning the specific lexical behavior of the remaining word list to the field of urban

planning. Also, range was prioritized over frequency to avoid biased due to topic-related words and word-length of journals. To meet the criteria of the range and frequency, vocabularies had to occur in at least half of the journals and 28.57 times per million word. Hence, to develop the field-specific word list, we included technical words that had the minimum range of 2 and minimum frequency of 29.

Results and Discussion

To identify the most frequent words of the field (Research Question 1), we followed the word selection criteria of Coxhead (2000), including range, frequency, and specialized occurrence. As Table 2 presents (Also see Appendix I for a complete set of core academic words), we found 507-word types prevalently appeared in Urban Planning research articles.

Table 2

Urban Planning Word List

Type	Range	Frequency
urban	4	3219
spatial	4	1557
density	4	787
traffic	4	642
climate	4	524
mobility	4	514
retail	4	476
household	4	425
zone	4	394
geography	4	350
census	4	336
rural	4	327
port	4	287
municipal	4	286
column	4	280
rental	4	276
upgrading	4	266
geographical	4	252
county	4	242

Table 3

Word Lists Coverage Across the Corpus

LEVEL	FILE	TOKEN	TOKEN %	TYPE	TYPE %
1	1_gsl_1st_1000.txt	869607	69.54	3469	10.7
2	2_gsl_2nd_1000.txt	58924	4.71	2119	6.53
3	3_awl_570.txt	157000	12.56	2578	7.95
4	Non-GSL-AWL	164922	13.19	24268	74.82

One of the efficient ways to evaluate the representativeness of a specific corpus and word list is the coverage, referring to the proportion of the total word count of a text or corpus (Kemp,2024). As delineated in Table 3, the General Service List (GSL) (West, 1953), encompassing 5588-word types, constitutes 928,531 tokens, which accounts for 74.25% of the entire corpus. The results indicate the pivotal role of GSL in facilitating the reading and comprehension of Urban Planning research articles, as it covers a substantial proportion of the corpus. While the first list of GSL(West,1953) provides a coverage of 69.54%, the second one constitutes 4.71% of the entire corpus. Additionally, AWL(Coxhead,2000) covers 12.56% of the whole corpus, outcovering the second list of GSL(West,1953) by almost 7.85%. This reinforces the speculation that it is not necessary to follow a static pattern in learning word lists, GSL before AWL, and other more technical word lists afterwards, in order to have a good grasp and understanding

705

of the academic texts (Valipouri and Nassaji,2013).

Additionally, the Academic Word List (AWL) (Coxhead, 2000) accounts for 2578-word types, comprising 157,000 tokens, covers 12.56% of the entire corpus. However, there are 24268-word types that are not included in either of these lists, encapsulating 164,922 tokens, which represent 13.19% of the entire corpus. These words form the UPAWL, which includes 557- word types, after refining the list. Considering that the Non-GSL-AWL covers 13.19% of the Urban Planning research articles, while AWL covered merely 12.56%, it can be inferred that Non-GSL-AWL, later to be the Urban Planning Academic Word List, would be more effective and efficient than AWL in terms of vocabulary coverage. Additionally, Non-GSL-AWL of our corpus provided more coverage than that of Coxhead's AWL. This phenomenon highlights the significant role of developing field-specific word lists, given the proposed notion that different lexical items have different behaviors across a wide range of areas in terms of meaning, frequency, and collocation (Hyland and Tse, 2007). Laufer (1999) argues that L2 readers, seeking to comprehend academic research articles successfully, should aim to familiarize themselves with approximately 95% (around 3,000 words) of the vocabulary present in such texts. The combination of GSL and AWL covered almost 87% of the corpus, indicating Urban Planning's academic community who might struggle when dealing with and comprehending such academic texts. Given the paramount significance of vocabulary acquisition for

academic achievement, researchers try to develop word lists to address students' needs in language learning with academic and specific purposes.

4.1 Technical Acronyms

Following the exclusion of non-academic terminologies and proper nouns, a list of acronyms was developed. As such, acronyms that were inconsistent with the conventions of technical academia were eliminated. As a result, only those acronyms that represented the academic connotations of Urban Planning were preserved and incorporated into the Urban Planning academic word list. Table 4 shows some of the identified technical acronyms of the field.

Table 4

Technical Acronyms of the Field

Acronym	Full Form
DCLG	Department for Communities and Local Government
OA	Opportunity Area
EPI	Energy Performance Index
GLA	Gross Leasable Area
CCP	Cities for Climatic Protection
EIA	Environmental Impact Assessment
PRS	Private Rented Sector
LSOA	Lower Layer Super Output Areas
GDP	Gross Domestic Product

Findings of this study provide pedagogical implications for EAP end-users. Given the significance of research-informed teaching (Richard et al., 2020), the immediate contribution of this study is for Urban Planning students and instructors to boost their discipline specific vocabularies to

communicate smoothly in the discourse community. This approach attempts to endow the instruction of language for specific and academic purposes with discipline-oriented evidence rather than the mere intuition of instructor about which techniques and strategies will work better and regard language teaching to be more accurate and effective. Such evidence-based approaches empower the instructors to focus on the students' language needs, which aids in the enhancement of the learning process in general and the understanding of specific language in particular contexts. More specifically, such research-driven teaching might expedite the consistent gap recently highlighted in the literature on the research-teaching nexus. Given the recent emergence of data-driven learning (DDL), instructors could use this discipline-specific word list and specialized corpus to adopt a cutting-edge approach to teaching (Sepehri, 2024). DDL represents a learner-centered methodology that facilitates the discovery of linguistic patterns and meanings. This approach encourages language learners to analyze extensive samples of authentic language usage, thereby enhancing their understanding of how language functions in various contexts (Perez-Paredes et al., 2019).

Second, for the field of EAP and ESP, this study is beneficial to materials developers who may incorporate the Urban Planning academic word list into the educational curricula, empowering a more specialized approach to teaching the Urban Planning area, given the high significance of language-focused learning, as proposed by Nation and Huntson (2018). Furthermore, not only may

the Urban Planning community find the Urban Planning academic word list a reliable reference for their desired purposes, but also active researchers of the target community may find it quite useful to have at their disposal to both write research articles and increase their chance to publish research articles in top-tier journals, leading to their professional and academic achievements.

Third, this study serves as a valuable resource for researchers and students. As Coxhead and Byrd (2007, p. 143) stated, "Students and teachers need to see vocabulary in the context of writing and not just reading." Therefore, by providing a large amount of reliable data on urban planning for the target audience, the researchers attempted to facilitate the process of learning vocabulary within an educational and research context. It is important to note that a list of vocabulary alone cannot fully meet students' needs, as vocabulary must be used in the relevant context to be effectively applicable. Ultimately, this academic vocabulary list in urban planning can be highly useful and practical for students intending to pursue this field at the university and academic settings as it aids students in engaging with specialized content in the field, such as textbooks, research articles, and other resources.

Conclusion

Aiming to address the lacuna of a specialized corpus of Urban Planning and a reliable word list, we carried out this study, both to identify core academic words of the field and determine the extent of assistance that GSL and AWL provide to the Urban Planning

707

community. Hence, we compiled and analyzed 100 empirical research articles that were published from 2013-2017 and 2018-2022, Applying word selection criteria (Coxhead,2000), and excluding irrelevant terms, irrelevant acronyms, and proper nouns, a discipline-specific word list of Urban Planning, consisting of 507-word types, was developed. Thus, this study provides a uniquely represented and globalized word list that has appeared in leading journals of urban planning, signifying the specific registers and nomenclatures that emerged in such dataset.

References

Alhasnawi, S. (2021). English as an academic lingua franca: Discourse hybridity and meaning multiplicity in an international Anglophone HE institution. Journal of English as a Lingua Franca, 10(1), 31-58. 10.1515/jelf-2021-2054

Anthony, L. (2021). AntWordProfiler (1.5.1w). Tokyo, Japan: Waseda University.

https://www.laurenceanthony.net/software/Antwordprofler

Barraza, A., Eckstein, G., Lambert, Z., Rogers, C., & Vincent, J. (2024). The applied linguistics

academic word list (ALAWL). Journal of Academic Language and Learning, 18(1), 21-48.

Brysbaert, M., & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of

current word frequency norms and the introduction of a new and improved word frequency measure for American English. Behavior Research Methods, 41(4), 977–

990. https://doi.org/10.3758/BRM.41.4.977.

Coxhead, A. (2000). A new academic word list. TESOL Quarterly, 34, 213–238. https://doi.org/10.2307/3587951

Coxhead, A., & Nation, P. (2001). The specialised vocabulary of English for academic

purposes. Research perspectives on English for academic purposes, 252-267.

Coxhead, A., & Byrd, P. (2007). Preparing writing teachers to teach the vocabulary and grammar

of academic prose. Journal of Second Language Writing, 16(3), 129–147. https://doi.org/10.1016/j.jslw.2007.07.002.

Farsani, M. A., Jamali, H. R., Beikmohammadi, M., Ghorbani, B. D., & Soleimani, L. (2021).

Methodological orientations, academic citations, and scientific collaboration in applied linguistics: What do research synthesis and bibliometrics indicate?. System, 100.

Farahani, Elham, Yazdani, Hooshang, Ahmadian, Moussa, Amerian, Majid (2020). An Investiagtion of Idioms in Applied Linguistics Research Articles: A Corpus-driven Study. Journal of Foreign Language Research, 10(2), 390-405.

Gholami Nejad, Razieh, Anani Sarab, Mohammad Reza. (2020). An investigation of the frequency of academic vocabulary in a body of academic textbooks of English language teaching. Journal of Foreign Language Research, 10(1), 206-222. https://doi.org/10.22059 / jflr.2020.289171.682

Hsu, W. (2013). Bridging the vocabulary gap for EFL medical undergraduates: The

establishment of a medical word list. Language Teaching Research, 17(4), 454- 484. https://doi.org/10.1177/136216881349412 1

Hyland, K., & Tse, P. (2007). Is there an "academic vocabulary"?. TESOL quarterly, 41(2), 235-

253. https://doi.org/10.2307/40264352

Kemp, J. (2024). How do I know this Law corpus is reliable and valid? Using a

representativeness argument for corpus validation. Applied Corpus Linguistics, 4(3), 1-10. https://doi.org/10.1016/j.acorp.2024.1000 99

Khani, R., & Tazik, K. (2013). Towards the development of an academic word list for applied

linguistics research articles. RELC journal, 44(2), 209-232.

Khany, R. and Kalantari, B. (2021). Accounting Academic Word List (AAWL): A Corpus-Based Study. Journal of Foreign Language Teaching and Translation Studies, 6(1), 35-58. doi: 10.22034/efl.2021.268643.1070

Kongnui, M. S., & Phoocharoensil, S. (2020). A corpus-based study of an English musicology word list. https://doi.org/10.14457/TU.the.2020.124 2

Kopchak, L., Zhumbei, M., Terletska, L., Konchovych, K., & Kanonik, N. (2022). The use of English as lingua franca to overcome language barriers and raise the level of education in modern conditions. Eduweb Magazine, 16(3), 121–133.

https://doi.org/10.46502/issn.1856-7576/2022.16.03.9

Korzin, A. S., Zhandarova, A. S., & Volkova, Y. A. (2023). Corpus-based approach to developing teaching materials for aerospace English. GEMA Online Journal of Language Studies, 23(3), 127-158

Laufer, B., & Nation, P. (1999). A vocabulary-size test of controlled productive ability. Language Testing, 16(1), 33–51. https://doi.org/10.1177/026553229901600103

Liu, J., & Han, L. (2015). A corpus-based environmental academic word list building and its validity test. English for Specific Purposes, 39, 1-11.

Martinez, I. A., Beck, S., & Panza, C. B. (2009). Academic vocabulary in agricultural research articles: A corpus-based study. English for Specific Purposes, 28(3), 183–198. https://doi.org/10.1016/j.esp.2009.04.003

Matinparsa, H., Xodabande, I., Ghafouri, M., & Atai, M. R. (2023). Academic vocabulary in applied linguistics research articles: A corpus-based replication study. Language Related Research, 14(5), 199-221.

Moini, R., & Islamizadeh, Z. (2016). Do we need discipline-specific academic word lists? Linguistics academic word list (LAWL). Teaching English as a Second Language Quarterly (Formerly Journal of Teaching Language Skills), 35(3), 65-90. https://doi.org/10.22099/jtls.2016.3901.

Milica Vuković-Stamatović. (2024). Creating and validating a corpus-based English academic word list for physics. Revista Española de Lingüística Aplicada/Spanish Journal of Applied Linguistics. https://doi.org/10.1075/resla.22041.vuk

Mudraya, O. (2006). Engineering English: A lexical frequency instructional model. English for Specific Purposes, 25(2), 235-256.

Nation, I. S. P. (2001). Learning vocabulary in another language. Cambridge University Press.

Nation, P., & Hunston, S. (2018). Learning vocabulary in another language. Cambridge University Press.

Özer, Mustafa & Akbas, Erdem. (2024). Assembling a justified list of academic words in veterinary medicine: The veterinary medicine academic word list (VMAWL). English for Specific Purposes. 74, 29-43. https://doi.org/10.1016/j.esp.2023.12.002.

Pérez-Paredes, Pascual & Ordoñana Guillamón, Carlos & Van de Vyver, Julie & Meurice, Alice & Aguado, Pilar & Conole, Grainne & Hernández, Purificación. (2019). Mobile data-driven language learning: Affordances and learners' perception. System. 84, 145-159. https://doi.org/10.1016/j.system.2019.06.009.

Plonsky, L. (2013). Study quality in SLA: An assessment of designs, analyses, and reporting practices in quantitative L2 research. Studies in Second Language Acquisition, 35, 655–687.

Plonsky, L. (2014). Study quality in quantitative L2 research (1990-2010): A methodological synthesis and call for reform. The Modern Language Journal, 98, 450–470.

Joseph- Richard, P., Almpanis, T., Wu, Q., & Jamil, M. G. (2021). Does research-informed teaching transform academic practice? Revealing a RIT mindset through impact analysis. British Educational Research Journal, 47(1), 226-245.

Roesler, David. (2021). When a bug is not a bug: An introduction to the computer science

academic vocabulary list. Journal of English for Academic Purposes. 54. http//doi.org/ 101044. 10.1016/j.jeap.2021.101044.

Sepehri, M. (2024). The Effects of data-driven learning on EFL learners' improvement in micro level skills of writing. Journal of Foreign Language Research, 14(3), 473-506. http//doi.org/ 10.22059/jflr.2024.379795.1143.

Swales, J. M. (1990). Genre analysis: English in academic and research setting. Cambridge University Press.

Valipouri, L., & Nassaji, H. (2013). A corpus-based study of academic vocabulary in chemistry research articles. Journal of English for Academic Purposes, 12(4), 248-263.

http//doi.org/10.1016/j.jeap.2013.07.001

Wang, J., Liang, S. L., & Ge, G. C. (2008). Establishment of a medical academic word list. English for Specific Purposes, 27(4), 442-458.

Ward, J. (2009). A basic engineering English word list for less proficient foundation engineering undergraduates. English for Specific Purposes, 28(3), 170–182. https://doi.org/10.1016/j.esp.2009.04.001

West, M. (1953). A general service list of English words, with semantic frequencies and a supplementary word-list for the writing of

popular science and technology. In Longman eBooks.

Wood, A. (2001). Research perspectives on English for academic purposes (J. Flowerdew & M. Peacock, Eds.). Cambridge University Press.

Xodabande, I., Atai, M. R., Hashemi, M. R., & Thompson, P. (2023). Developing and validating a mid-frequency word list for chemistry: a corpus-based approach using big data. Asian-Pacific Journal of Second and Foreign Language Education, 8(1). https://doi.org/10.1186/s40862-023-00205-5

Xodabande, I., & Xodabande, N. (2020). Academic vocabulary in psychology research articles: A corpus-based study. MEXTESOL Journal, 44(3).https://doi.org/10.1186/s40862-023-00205-5

Xodabande, I., & Xodabande, N. (2020). Academic vocabulary in psychology research articles: A corpus-based study. *MEXTESOL Journal*, *44*(3).