

Constructing the LERWL: A Corpus-Based Approach to Identifying Technical Vocabulary in Language Education Research

SUPAKORN PHOOCHAROENSIL

Language Institute, Thammasat University, Thailand

Author email: supakorn.p@litu.tu.ac.th

Article information	Abstract
Article history: Received: 16 Mar 2026 Last revised: 5 Jul 2026 Accepted: 6 Jul 2026 Available online: 10 Jul 2026 Keywords: Word list Language education research Corpus-based study Technical vocabulary ESP	<i>Technical vocabulary plays a crucial role in ESP learners' lexical development once general service and academic vocabulary has been established. The present study constructed the Language Education Research Corpus (LERC), an 8,646,901-word corpus compiled from the top 10 Quartile 1 (Q1) Scopus-indexed journals in Language Education. On the basis of this corpus, the Language Education Research Word List (LERWL) was developed through four systematic procedures. High-frequency items were first identified, after which a range criterion was applied to retain words occurring in at least 50% of the target journals. Lexical profiling was subsequently conducted to isolate discipline-specific vocabulary, excluding items listed in the GSL and the AWL. In addition to these quantitative procedures, expert judgment was employed as a qualitative validation measure to ensure disciplinary relevance. The resulting LERWL was then evaluated against (a) the LERC, achieving 5.28% coverage, and (b) an independent corpus of approximately one million words from the same field, achieving 4.75% coverage. These findings indicate that the LERWL provides a substantial and pedagogically meaningful level of lexical coverage within the discipline.</i>

INTRODUCTION

Vocabulary is a fundamental component of language use and thus requires significant emphasis in language teaching (Szudarski, 2018). Nation (2022) proposed the principle of the four strands, which emphasizes a well-balanced course comprising four equal strands: meaning-focused input, meaning-focused output, language-focused learning, and fluency development. In the strand of learning from comprehensible meaning-focused input, learners are expected to acquire new vocabulary in the second language (L2) through listening and reading, where the focus is on the information they are exposed to. In the strand of meaning-focused output, learners' L2 lexical knowledge improves through speaking and writing activities, during which attention is given to the information they are attempting to convey. Thirdly, the strand of language-focused learning, also known as form-focused instruction, implies that an L2 lesson requires direct instruction of vocabulary. Finally, the strand of fluency development does not involve the learning of new vocabulary; instead, learners become increasingly fluent in the use of the lexicon they already know.

Once learners have mastered the 2,000–3,000 commonly used words in general English, their next objective should be to acquire specialized vocabulary. This includes academic terminology that spans across different fields, followed by a focus on the specific jargon of their chosen discipline (Nation, 2022). This technical vocabulary, which is closely tied to a particular subject area, is invaluable for English learners aiming to use English for Specific Purposes (ESP), such as comprehending discipline-specific texts, composing technical reports, or participating in subject-specific seminars. Emphasizing technical terms is essential in ESP courses, as they serve as a solid foundation for acquiring ESP vocabulary (Nation, 2016). It is important not to overlook unfamiliar technical terms while reading, as they are integral to understanding the text's subject matter. Guessing their meanings solely from context can be challenging unless learners possess a solid background in the subject area. Simply put, as Nation (2016) suggests, some specialist areas “have a very large proportion of words known only by specialists, making them difficult areas to study” (p. 7). Recent studies have focused on compiling technical word lists to determine their scope and to explore how learners in specialized fields can enhance their grasp of technical vocabulary using these lists (e.g., Arunvong Na Ayutthaya et al., 2022; Kruawong & Phoocharoensil, 2020; Laosrirattanachai & Ruangjaroon, 2020, 2021a, 2021b; Rungrueang et al., 2022).

Language Education (LE), which encompasses the teaching and learning of languages, holds significant importance within the realm of social sciences. Fields related to LE include English Language Teaching (ELT), Teaching English to Speakers of Other Languages (TESOL), Teaching English as a Foreign Language (TEFL), and Teaching English as a Second Language (TESL) (Hall, 2016). Many individuals, including undergraduate and graduate students as well as early-career researchers in LE, may find it challenging to engage with textbooks or research articles in this field. LE covers a broad spectrum of research fields, and there have been notable advancements in research trends within this discipline. Initially, LE research primarily focused on linguistic analysis, case studies, and statistical studies during the early 1980s. However, over the following decades, both quantitative and qualitative research methods have expanded into various forms. Quantitative methods in LE now include descriptive studies, exploratory studies, quasi-experimental studies, and experimental studies, while qualitative methods cover diverse areas such as action research, conversation analysis, discourse analysis, ethnographies, interviews, introspection methods, narrative inquiries, open-ended questionnaires, and teacher inquiries (Mackey & Gass, 2023). Familiarizing LE students and researchers with technical terminology through formal instruction could prove advantageous, as creating a specialized word list for LE has the potential to enhance the materials used by LE teachers, such as textbooks, and improve students' comprehension of texts within this discipline.

This study aims at developing the Language Education Research Word List (LERWL) based on the Language Education Research Corpus (LERC). This study also reports the creation of the LERC based on research articles published in Quartile 1 (Q1) Scopus-indexed journals in the field of LE, as listed in the SCImago Journal Rank as of April 2024. The search was confined to journal titles directly related to language teaching, learning, and testing, with journals emphasizing theoretical linguistics or lacking relevance to language education being excluded. This study excludes journals with non-English titles, as they are likely to publish

articles in languages other than English. Additionally, only articles from journals that are accessible to the researcher were considered. The Top 10 journals, all ranked in Q1 of Scopus, were selected, and articles published between 2019 and 2023 were included in the analysis.

LITERATURE REVIEW

Vocabulary categories in language learning and teaching

Second-language (L2) learners are required to grasp considerable vocabulary to effectively understand and use the target language, be it spoken or written. Through an investigation into different kinds of texts using 1,000-word family lists based on the British National Corpus (BNC), Nation (2006) reveals that it is necessary for learners to know approximately 3,000–4,000 word families to achieve 95% text coverage, which suffices for understanding spoken narrative texts (van Zeeland & Schmitt, 2013). To achieve 98% coverage, learners typically need to know between 6,000–9,000 word families. It is suggested that vocabulary acquisition follows a frequency-based pattern, where the most common words are learned first, followed by less frequent ones. This implies that learners tend to acquire the first 1,000 words before moving on to the next set, and so forth. In a nutshell, the frequency of L2 vocabulary plays a vital role in determining the priority of words to be introduced to learners (Nation, 2022).

Pedagogically speaking, high-frequency words should receive the greatest attention in L2 vocabulary instruction due to the fact that learners who have acquired high-frequency words have a greater chance “...to enrich their knowledge through later meetings with the words, and will have the greatest opportunities to produce what they know” (Nation, 2016, p. 5). The classic high-frequency word list is Michael West’s (1953) General Service List (GSL), which includes mainly content words making up around 2,000 word families. Bearing in mind that high-frequency words are useful to students, enabling them to get the best return for their vocabulary learning effort, teachers tend to prioritize these common words in their lessons. Following the high-frequency words are the mid-frequency ones, e.g., *aired*, *pioneering*, and *pastoral*. There are 6,000 to 7,000 mid-frequency words in English, ranging from the third 1,000 words to the ninth 1,000 words (Nation, 2022). Nation (2022) and Schmitt and Schmitt (2012) emphasize the importance of incorporating the high-frequency and mid-frequency words in ELT curricula as they both constitute the amount of vocabulary required for effective English use without the need for outside support. Teachers aware of such a benefit should focus on these two levels of vocabulary and thus set learning goals for students (Nation & Webb, 2011).

Vocabulary beyond the first 9,000 English words is regarded as the low-frequency words, e.g., *eponymous*, *gibbous*, and *bifurcate* (Nation, 2022). There are ten of thousands of such low-frequency words in English, although considered a very large group of vocabulary, and they cover only a small proportion of any text (Nation, 2022). It is interesting to note that many low-frequency words are proper names, e.g., *Carl*, *Johnson*, and *Ohio*, which may occur frequently in certain texts, e.g., characters’ names in novels, rather than general

texts, and they form a huge proportion of the word types, i.e., approximately 50%, in any large corpus. However, it is noteworthy that low-frequency words include technical terms for a variety of disciplines and words that learners hardly encounter in everyday life. Needs for technical vocabulary vary from one person to another. Due to interests and specializations, one person's technical vocabulary can be seen as another person's low-frequency words. Doctors, for instance, are undoubtedly familiar with and find medical terms like *biopsy*, *edema*, and *defibrillator* important in their field, while accountants seem to consider them words with very low frequency.

Another lexical category is specialized vocabulary associated with specific text types, including academic and technical words. Academic words represent the language commonly used in academic contexts. One of the best-known lists containing common academic vocabulary is Averil Coxhead's (2000) Academic Word List (AWL). The AWL consists of 570 word families not included in the most frequent 2,000 English words but commonly found in a wide range of academic texts. It comprises words from the third 1,000 to the seventh 1,000 words. More precisely, the AWL covers four main disciplines: humanities, law, science, and commerce, making it essential for learners of English for Academic Purposes (EAP).

In contrast to the AWL, Gardner and Davies (2014) developed the Academic Vocabulary List (AVL). The AVL comprises academic words extracted from the Corpus of Contemporary American English (COCA). Its objective was to present an updated compilation of essential or 'core' academic vocabulary, excluding both common high-frequency words and subject-specific technical terminology. To ensure the inclusion of exclusively academic vocabulary, Gardner and Davies employed a criterion based on the frequency of words in academic versus non-academic contexts. Words were included only if they appeared at least 1.5 times more frequently in academic texts compared to non-academic ones, representing a 50% higher occurrence rate in the academic corpus. This approach is different from the method used to develop the AWL, which involved the exclusion of words from the GSL.

Unlike the AWL, based on word families, the AVL was constructed using lemmas, encompassing words and their inflected forms. For instance, the verb *provide* has inflected forms *provided*, *provides*, and *providing*. This distinction addresses another critique of the AWL, which arises from the fact that different members of a word family may not share identical core meanings, and some forms within the family may have extremely low frequencies. Grasping these nuances in meaning and identifying the more commonly used forms within a word family pose significant challenges for English learners. Another distinction between the AVL and AWL lies in their approach to excluding general words. While the AVL accomplished this by considering the ratio of words in academic versus non-academic texts, the AWL simply omitted words from the outdated GSL. Additionally, there is a difference in the size of the corpora utilized. The AVL was developed from a 120-million-word subset of the 425-million-word COCA. By contrast, the AWL relied on a much smaller corpus, comprising only 3.5 million words.

Aside from academic words, specialized words also include technical vocabulary, i.e., words that are common in a topic area but are infrequent in other subjects. In other words, technical vocabulary is composed of words closely related to the content of a specific discipline (Nation,

2016). Technical vocabulary can be derived from any of the aforementioned three frequency levels of vocabulary. Some high-frequency words can be technical words in some fields. As illustrated by Nation (2022), *arm*, *leg*, and *neck*, all of which are very frequent words in general English, are also labeled technical words in anatomy. Meanwhile, some mid-frequency words can have technical meanings in some disciplines. Words with low frequency in one discipline may also be technical words in another discipline.

Technical words can be identified in various ways. One of the most popular approaches is to seek expert advice to determine the relevance of a word to the subject matter of the discipline. Some researchers also consult technical dictionaries to help identify technical vocabulary in a particular field (Nation, 2022). However, reliance on such dictionaries is not without problems. Major criticisms against their use include a lack of clear criteria for determining word inclusion, and the fact that a dictionary reflects the output of a single lexicographer rather than a consensus among experts in the field (Nation, 2022). Another reliable method for distinguishing technical vocabulary from general vocabulary is through corpus comparison. Chung (2003) compared word frequencies in a general corpus, comprising various texts such as novels, academic texts, newspapers, and magazines, with those in a specialized technical corpus. Words occurring at least 50 times more frequently in the technical corpus, thus unique to it, are categorized as technical words. The resulting technical word list was found to be around 90% consistent with a specialist's judgment. Developing such a technical word list is highly beneficial for both teachers and learners in an ESP context, as it enables access to authentic technical vocabulary common in a particular discipline. In the following section, the theoretical background in word list compilation and related studies is reviewed.

Word lists

A word list is referred to as a list of all the words in a corpus, normally extracted through some automatic methods and such a list can then be sorted alphabetically or by frequency (Flowerdew, 2012, p. 324). A word list is usually generated using a software program e.g., *AntWordProfiler* (Anthony, 2023) and *VocabProfile* (Cobb, 2021). Two main types of word lists are classified, namely frequency-based word lists and specialized word lists, both of which are beneficial to L2 learners (Nation, 2016).

Researchers have been developing vocabulary lists through decades with the aim of providing core English vocabulary for ELT practitioners. However, Nation (2016) notes that some early lists, e.g., Basic English comprising 850 words, created by Ogden (1930) and Richards (1943), were primarily based on the authors' subjective judgement rather than frequency. Among the major limitations of Basic English, in addition to the use of intuition-based approach, are the inclusion of polysemous words, e.g., *head* and *bear*, and a lack of direction for learners to move toward more advanced vocabulary (Szudarski, 2018). In short, it seems very challenging to create a core word list that is both valid and reliable.

With regard to frequency-based word lists, Nation (2022), as mentioned earlier, differentiates vocabulary into three groups based on word-family frequency lists: high-frequency words, mid-frequency words, and low-frequency words. Nation (2022) also highlights the significance of frequency in vocabulary lesson preparation and suggests that teachers focus more on

word frequency. Frequency analysis enables researchers and teachers to have a clearer understanding and description of vocabulary use. Knowledge of word frequency can guide teachers and curriculum developers in deciding which vocabulary to focus on during language instruction and determining the vocabulary included in textbooks (Szudarski, 2018, p. 66). A traditional frequency-based list is West's (1953) GSL, a useful list which comprises 2,000 word families and information on their frequency, different meaning-senses, grammatical classes to which they belong and lexical phrases in which they usually occur. The GSL has been challenged due to the researcher's subjective judgments in the selection of the target words and its obsolescence, compared to next-generation word lists, such as the New General Service List or new-GSL (Brezina & Gablasova, 2015), containing 2,494 lemmas identified on the basis of the frequency analysis of 12 billion words from four corpora.

Apart from the aforementioned frequency-based word lists, developing specialized word lists is also worthwhile for ELT. In this study, the definition of specialized word lists is equivalent to that of technical word lists (Liu & Han, 2015). Nation (2016) highlights several benefits of such word lists. First, studies on technical vocabulary can contribute to vocabulary size research. It is important to note that while some technical words are accessible through general knowledge, e.g., *arm*, *lungs*, and *penicillin*, some are only available through specialized knowledge, e.g., *costal*, *xiphoid*, and *haemoglobin*, which are all considered technical words in the field of medicine. The second benefit of specialized word lists is that they aid in understanding the magnitude of the vocabulary learning task learners face when beginning their studies in technical areas. Third, research on technical vocabulary also enables teachers to develop subject-specific materials for EAP and ESP courses.

Criteria for word list development

In the development of specialized word lists, it is customary for certain criteria to be fulfilled: frequency, range, keyness, lexical profiling, and expert judgement.

Frequency

When creating corpus-based word lists, researchers primarily prioritize frequency, namely how often a word appears in a corpus (Nation, 2016). In the field of ELT, it is prudent to focus more on words with higher frequency rather than those used less frequently (Webb & Nation, 2017). However, frequency alone should not be the sole criterion for selecting useful words for teaching L2 vocabulary. Among the most frequently occurring words are function words such as articles, e.g., *the*, pronouns, e.g., *I* and *you*, conjunctions, e.g., *and* (O'Keefe et al., 2007), which do not reflect discipline-specific vocabulary. Another weakness of using only frequency in word list compilation is that a word identified as frequent in one particular text may rarely appear in others. For example, as Nation (2016) points out, the infrequent word *profitboss*, despite its frequency of 425, occurs in only one text in the BNC.

Range

Range shows the number of different texts or sub-corpora in which a word occurs. Of all the statistical criteria for the vocabulary inclusion in a word list, range holds paramount

importance (Nation, 2016). For the validity of range figures, each sub-corpus should be of equal size and internally coherent, necessitating similarity in text types across sub-corpora. Vocabulary experts such as Coxhead (2000), consistent with Nation (2016), advocate considering range to create a balanced word list. However, a limitation of using range is that some words may appear in numerous texts but with low frequency. In a corpus divided into several sub-corpora, a word may be present in all sub-corpora, resulting in a high range figure, yet it may occur only once or twice in some of them. In essence, a high range figure does not guarantee high frequency.

Keyness

Word list developers can also perform keyword analysis, which goes beyond simple frequency counts and involves comparing two frequency lists (Tangpijaikul, 2014). In keyword analysis, the ranking of a word with exceptionally high frequency in a given corpus is compared with that in a reference corpus, such as the BNC or the COCA, which is larger or of equal size to the corpus being examined. A word will be designated as 'key' if it occurs "...in a corpus statistically significantly more frequently when compared with its occurrences in a reference corpus..." (Flowerdew, 2012, p. 321).

Lexical profiling

To determine technical vocabulary in a specific discipline is not at all an easy task. While some researchers compare words in their compiled corpora with those in technical dictionaries, assuming that words absent from the dictionaries are non-technical, this approach has significant drawbacks. Such dictionaries rely on the subjectivity of specific writers and often lack comprehensiveness, failing to include every technical term in a given subject (Nation, 2022). Consequently, corpus-based word list researchers often turn to a more reliable alternative known as lexical profiling. This method involves filtering out words found in the GSL and the AWL, which serve as typical reference corpora, though other word lists may also be used as reference corpora. It is worth noting that, despite its limitations, the AWL is widely used as a reference corpus in many software programs (e.g., *AntWordProfiler*), as it remains one of the most well-known academic vocabulary lists. It has also been adopted by several major learner's dictionaries to label academic vocabulary. Lexical profiling is advantageous as it removes words irrelevant to the field under investigation (Laosrirattanachai & Laosrirattanachai, 2021). However, a critical issue with this criterion is that some words related to the specific fields being examined may be excluded if they appear in the GSL, AWL, or other reference word lists used in the lexical profiling process (Namwong et al., 2022).

Expert judgement

It is advisable to utilize this criterion in the final stage of constructing a specialized word list. Chung and Nation (2003) argue that while some words, such as *security*, have a general English meaning of 'the state of being safe', they can acquire a technical meaning in business, such as 'some real estate asset needed for a bank loan'. The meaning of a word seems to depend on the contexts in which it is used (Tangpijaikul, 2014). Seeking advice from experts

involves using a questionnaire listing specific words along with a 4-category rating scale (Chung & Nation, 2003):

Scale 1 identifies a word whose meaning is not relevant to the specific discipline.

Scale 2 identifies a word whose meaning is little relevant to the specific discipline.

Scale 3 identifies a word whose meaning is very relevant to the specific discipline.

Scale 4 identifies a word that is only used in the specific discipline.

The use of a rating scale approach benefits the construction of specialized word lists by assisting researchers in identifying more relevant words, specifically those rated between 3 and 4 (Namwong et al., 2022), and in retrieving related technical words that were previously excluded through lexical profiling.

Over the past ten years, 66% of the published word list studies have concentrated on technical vocabulary (Dang & Webb, 2025). A number of technical word lists have been constructed to the benefit of ESP instruction. Such ESP lists cover a wide range of disciplines, such as science (Coxhead & Hirsh, 2007; It-ngam & Phoocharoensil, 2019), chemistry (Valipouri & Nassaji, 2013; Xodabande et al., 2023), zoology (Kruawong & Phoocharoensil, 2020), engineering (Hsu, 2014; Todd, 2017; Ward, 2009), marine engineering (e.g., Đurović, 2021), nursing (Yang, 2015), veterinary medicine (Özer & Akbaş, 2024), COVID-19 (e.g., Saed et al., 2022), business (Konstantakis, 2007; Tangpijaikul, 2014), hospitality (Laosrirattanachai & Ruangjaroon, 2020, 2021), food service (Rungrueang et al., 2022), beverage service (Arunvong Na Ayutthaya, 2022), Thai tourism (Laosrirattanachai & Laosrirattanachai, 2021), agriculture (Martinez et al., 2009; Muñoz, 2015), geography (Drašler & Kavalir, 2026), political sciences (Nevisi et al., 2023), and applied linguistics (Khani & Tazik, 2013). While these existing ESP word lists have made significant strides in catering to specific domains, there remains a lack of lists on language education research. It is worth noting that language education is now not at all equal to applied linguistics. As Wei et al. (2024) remark, “[u]ntil at least the 1980s, applied linguistics was most closely associated with the problems and puzzles surrounding language pedagogy, learning, and acquisition”. Currently, applied linguists have new opportunities to explore emerging areas of research, for instance, translanguaging or translingual practices, interculturality, intersections between translation and politics and the law, bi- and multilingualism, media discourse, forensic linguistics, sign language, language planning and language policy, and family language policy, etc. (Wei et al., pp. 1–2). Given that language education and applied linguistics are not synonymous, reliance on word lists in Applied Linguistics (e.g., Khani & Tazik, 2013) may not directly address the needs of LE researchers and students. Therefore, creating a word list in language education is necessary and is regarded as beneficial for both learners and researchers in the fields of language learning and teaching, including ELT, TESOL, TEFL, and TESL.

Due to the aforementioned need for a specialized word list in LE, the researcher compiled a corpus, referred to as the LERC, comprising 8,646,901 words. The corpus was constructed from LE research articles published between 2019 and 2023 in Q1 Scopus-indexed journals. Based on the LERC, an LE vocabulary list was developed using frequency, range, lexical profiling, and expert opinion. The research methodology of word list construction is elucidated in the ensuing section.

RESEARCH METHODOLOGY

Compilation of the corpus

The specialized corpus, known as the LERC, was created based on research articles published in Q1 Scopus-indexed journals in the field of LE, as listed in the SCImago Journal Rank as of April 2024, when the website's latest information was updated. The SCImago Journal Rank is a measure of scientific journals' prestige and impact, developed by SCImago Lab. To display the journal ranks, the subject area *Linguistics and Language* were selected. Only journal titles relevant to language teaching, learning, and testing were targeted in the search. Journals focusing on theoretical linguistics or lacking implications for language education, e.g., *Neurobiology of Language*, *Russian Journal of Linguistics*, *Transactions of the Association for Computational Linguistics*, *Discourse*, and *Journal of Specialised Translation*, were excluded. Journals with non-English titles are not within the scope of this study, as they potentially publish non-English articles. To ensure the accessibility of journal articles for this study, only research articles available through databases accessible via the researcher's institute's subscription were downloaded. The top 10 journals, all of which are indexed in Q1 of Scopus, have been identified, and articles in all of the volumes published between 2019 and 2023 were collected. The finalized journal titles are presented in Table 1 below. As this study focuses on LE research, only research articles were included, while other types of publications (e.g., review articles) were excluded. It should be noted that, during the data preparation stage, textual content was extracted from all sections of the retrieved articles, from the abstracts through to the conclusions, with references and appendices excluded.

Table 1
The Q1 journal titles in language education

No.	Journal title	Number of articles	Word count
1	Studies in Second Language Learning and Teaching	126	917,505
2	Computer Assisted Language Learning	63	473,391
3	Journal of Writing Research	73	713,696
4	Language Learning and Technology	115	838,144
5	Language Testing in Asia	166	1,285,324
6	Iranian Journal of Language Teaching Research	107	727,007
7	Asian-Pacific Journal of Second and Foreign Language Education	128	951,780
8	Ampersand	103	758,666
9	TESL-EJ	210	1,500,806
10	ESP Today	70	480,582
	Total	1,161	8,646,901

Research instruments

In this study, two instruments were utilized. Firstly, the software program *AntWordProfiler* (Anthony, 2023) was employed to develop the LERWL. This software program is a linguistic tool designed for corpus analysis and the generation of various linguistic statistics and lexical information. In simpler terms, it provides information on the vocabulary used in a text, including word frequencies, ranges, word types, and measures of lexical diversity. *AntWordProfiler*

stands out as a modern vocabulary analysis program capable of performing lexical profiling, unlike other programs such as *VocabProfile* or *The Range program*. The second instrument was a questionnaire consisting of an initial word list created through frequency, range, and lexical profiling. In line with Laosrirattanachai and Ruangjaroon (2021a), this questionnaire was distributed to five experts in language education to determine the technical words in this particular discipline that should be included in the final list.

Data analysis

In the development of the LERWL, four criteria were applied: frequency, range, lexical profiling, and expert judgement, respectively.

Step 1: Frequency—*AntWordProfiler* was utilized to calculate the frequency of words in the LERC. According to Nation (2005), there is no strict criterion for defining a high-frequency cut-off point. This study therefore adopts Coxhead's (2000) method for frequency calculation, requiring a word to meet the frequency criterion if it appears at least 100 times in a corpus of 3,500,000 tokens. In the present study, the corpus size is 8,646,901 words; thus, the size of the LERC was compared with Coxhead's using the equation below:

$$\frac{100}{3,500,000} = \frac{X}{8,646,901}$$

$$X = 247.05 \text{ or } 247$$

As per the equation provided, a word must occur 247 times or more to qualify as frequent and be selected.

Step 2: Range—The range hinges on the total number of available sources. As there is no fixed rule for determining the range value of words in a word list (Nation, 2016), this study follows Coxhead's (2000) approach. In Coxhead's investigation, the academic corpus was gathered from 28 diverse sources, with specialized words required to occur in at least 15 sources, constituting roughly 50 percent of all sources, to fulfill the range criterion. For this reason, a technical word that was found in at least 5 out of 10 selected journals was extracted.

Step 3: Lexical Profiling—After applying the first two criteria to select technical words from the LERC, the researcher utilized *AntWordProfiler* to remove irrelevant words. This software can filter out words that also appear in the GSL and the AWL, ensuring that only technical terms relevant to language education remain after screening. Additionally, the Function Word List (FWL) (Nation, 2016), which includes function words like auxiliaries, modals, prepositions, articles, conjunctions, and pronouns, was uploaded into *AntWordProfiler* to eliminate function words from the specialized list. Lexical profiling enables the creation of targeted and effective technical word lists that cater to the needs of ESP learners or researchers in language education.

Step 4: Expert Judgment—The vocabulary list that passed the first three screening criteria was included in a questionnaire presented to five experts in language education, all holding

Ph.D. in language education or related fields. Chung and Nation's (2004) 4-category rating scale checkbox was employed to construct the questionnaire. A score of 1 denotes a word whose meaning is irrelevant to the specific discipline, while a score of 2 indicates a word whose meaning is somewhat relevant to the specific discipline. A score of 3 identifies a word highly relevant to the specific discipline, and a score of 4 indicates a word exclusively used in the specific discipline. Based on the intuitive judgment of the experts, any words receiving a score of 3 or 4 from at least three experts were included in the LERWL.

The words in the LERWL are presented as lemmas, that is, the canonical or base forms that represent all inflectional variants of a word. For example, the lemma *TRANSCRIBE* encompasses all its verb forms, including *transcribe*, *transcribes*, *transcribing*, and *transcribed*. Lemmas were used instead of word families because not all members of a given word family were identified in the corpus. For instance, while the adjective *phonological* was included, related forms such as the nouns *phonology* and *phonologist* and the adverb *phonologically* were not, likely because they did not meet the frequency or range criteria. This suggests that *phonological*, in contrast to other related forms, is more strongly associated with the field of language education.

Step 5: Validation of the LERWL's coverage—After the full version of the LERWL was developed, its coverage was evaluated to ensure that it adequately represents technical vocabulary in the discipline. According to Chung and Nation (2003) and Hyland and Tse (2007), the effectiveness of a technical word list should be assessed using an external corpus rather than the one from which it was derived, with an expected coverage of approximately 5% of the text. Accordingly, the LERWL was tested against a 1,162,451-word corpus compiled from articles published in the same ten journals in 2018, a year that falls outside the scope of the main study.

FINDINGS

After the LERC was created, the tokens of 8,646,901, i.e., the total number of words in the corpus, were changed into the type form, i.e., the number of different words. The resulting types contain 217,497 words. Then the four aforementioned criteria were applied to obtain the LERWL.

The frequency criterion

In the development of the LERWL, the initial criterion applied to the 217,497 identified word types was frequency. Specifically, the software AntWordProfiler was utilized to extract word types exhibiting a frequency of 247 or above, based on a total corpus frequency of 8,646,901. This thresholding procedure aligns with the methodology established by Coxhead (2000). Application of the frequency criterion resulted in the selection of 3,262 words. As expected, function words dominated the highest frequency ranks; examples include *the*, *of*, *an*, *in*, *to*, *a*, *that*, *for*, *as*, and *and*. However, the Top-40 list also contained significant content words, all of which appear to be relevant to the specific domain of study. These domain-specific content words encompassed terms such as *language*, *linguistic*, *English*, *translation*, *writing*, and *vocabulary*.

The range criterion

The 3,262 words initially identified via the frequency criterion were subsequently analyzed using the range criterion, which required the words to occur in a minimum of five out of the ten target journals. A substantial majority of the words, 3,232, ultimately satisfied both the frequency and range criteria. This outcome suggests that the high-frequency vocabulary within the corpus also exhibits a high degree of dispersion across the selected journals. Specifically, it is noteworthy that 2,967 words, e.g., *comprehension*, *feedback*, *identity*, *language*, *listening*, *vocabulary*, were found to occur in all ten journals, while only 30 words failed to meet the specified range requirement and were consequently removed.

The lexical profiling criterion

The words that met the established frequency and range criteria were subsequently subjected to lexical profiling via *AntWordProfiler*. As shown in Table 2, of the 3,232 words satisfying these criteria, 1,738 words (representing 53.77%) were identified in the GSL, and 823 words (representing 25.46%) were identified in the AWL. Following lexical profiling, the words belonging to the GSL and the AWL were removed, resulting in a final pool of 671 words, which were then prepared for expert review and validation; examples of the remaining words include *collocation*, *descriptor*, *fluency*, *genre*, *grammatical*, *intelligibility*, *interlocutor*, *literacy*, *metacognitive*, and *multimodal*.

Table 2
Lexical profile of candidate words

Profile	Number	%
1 st 1,000 GSL	1,404	43.44
2 nd 1,000 GSL	334	10.33
AWL	823	25.46
The rest	671	20.76
Total	3,232	100.00

The expert judgment criterion

The researcher deleted abbreviations and scholars' names, as they do not represent the discipline, from the 671 words filtered through the first three criteria. Thus, 511 words were included in the questionnaire which was sent out to five experts in language education, all of whom hold a Ph.D. in ELT or TESOL to select the words that are relevant to research in language education. The words assigned a score of 3 or 4 by the minimum of 3 experts were placed in the final list. Examples of words receiving a score of 4 by all the experts, which means they were all considered to be very relevant to language education, were *adjective*, *bilingual*, *comprehension*, *grammatical*, and *linguistic*. The vocabulary that satisfied the expert judgment criterion totaled 342 words.

Finalizing the LERWL and its coverage

The present study aimed to present the vocabulary in the word list in lemma form. For example, different forms such as *align*, *aligns*, *aligned*, and *aligning* were combined under the single lemma ALIGN. After the LERWL was created, the researcher used AntWordProfiler to check its coverage against the LERC and found that it covered 5.28%. This finding is consistent with Laosrirattanachai & Ruangjaroon (2020), who suggest that a technical word list should cover approximately 5% of a text. To further validate the LERWL's coverage, a smaller corpus of 1,162,451 words was compiled from articles published in 2018 in the same group of target journals. It is recommended that a word list be tested against external corpora (Brezina & Gablasova, 2015). The LERWL covered 4.75% of this corpus, which is considered sufficiently high to be representative of the discipline.

CONCLUSION AND DISCUSSION

The main objective of this study was to develop a word list that would facilitate language educators and language education researchers. The researcher began the project by creating the LERC, comprising 8,646,901 words drawn from the top 10 Q1 journals indexed in Scopus and published between 2019 and 2023. Based on this corpus, four major criteria, i.e., frequency, range, lexical profiling, and expert validation, were employed to select the technical words. The combination of quantitative criteria (i.e., frequency, range, and lexical profiling) and a qualitative criterion (i.e., expert judgment) contributed to a more robust final word list.

Specifically, this study established a minimum frequency threshold of 247 occurrences and a minimum range of five out of the ten journals to identify technical words in the LERC. Words that met both quantitative criteria then underwent lexical profiling to exclude content words appearing in the GSL and the AWL, that is, words not specific to the LER domain. The resulting LER word list of 671 words was subsequently sent to five experts in the field for validation. Reliance on expert judgment as a qualitative criterion added further rigor to the study.

The effectiveness of the LERWL of 342 words was validated through two distinct stages: an analysis using the LERC and a subsequent evaluation against a smaller, newly compiled corpus. The results demonstrated the high lexical coverage of the LERWL within the field of language education research.

As a technical word list, the LERWL has several advantages in language teaching and learning. First, as stated by Nation (2016), by referring to the LERWL, from the instructional perspective, teachers can set clear learning goals for students, particularly at the graduate level, majoring in language education, as well as for students as emerging researchers in the field. By consulting this list, teachers can more accurately identify target vocabulary during lesson planning, moving beyond mere intuition when selecting technical terms. For example, teachers preparing students to read research articles in vocabulary learning and teaching can refer to the LERWL to introduce key technical terms such as *lexis*, *lexical*, *collocate*, *collocation*,

gloss, idiom, and VLS (Vocabulary Learning Strategy). Since the LERWL is derived from four systematic criteria, it empowers curriculum developers to design ESP materials grounded in authentic, domain-specific vocabulary (It-ngam & Phoocharoensil, 2019).

Second, from the learners' perspective, students of (ESP), who are typically at an advanced level, must master technical vocabulary after acquiring general proficiency. Consequently, language education researchers already familiar with the GSL or AWL must bridge the gap by mastering the LERWL to fully comprehend discipline-specific texts. To illustrate, ELT-major students enrolled in a course on the linguistic foundations of ELT may find terms such as *classifier, consonant, entail, discourse, and phonetics* particularly useful, as familiarity with these terms can facilitate comprehension of textbooks and research articles. Proficiency in this specialized lexicon enables learners to navigate communication challenges within the field more effectively (Dang & Webb, 2025). Furthermore, access to these curated terms fosters learner autonomy, significantly enhancing their ability to decode and analyze complex research articles in this specialized area (Saed et al., 2022).

Third, the LERWL also benefits testing. ESP teachers are encouraged to design corpus-based reading and listening tests using the prepared technical vocabulary from word lists such as the LERWL. Selecting technical words from the LERWL for inclusion in a text or passage can help teachers assess not only English proficiency but also field-specific content knowledge (Kruawong & Phoocharoensil, 2024). This is particularly beneficial for students in English programs who are taught through Content and Language Integrated Learning (CLIL), an educational approach in which students learn non-linguistic subjects through a second, foreign, or additional language (Coyle et al., 2010). Test developers can also draw on the LERWL as a reliable source of technical vocabulary within the discipline. It is considered a better source of lexical items than specialized dictionaries, whose word selection criteria are often unclear (Nation, 2016).

LIMITATIONS AND RECOMMENDATIONS FOR FUTURE RESEARCH

It is important to acknowledge some limitations of this study. First, the final word list was developed based on four criteria: frequency, range, lexical profiling, and expert validation. Future studies could incorporate keyword analysis, a more advanced and complex criterion, to filter technical vocabulary, which may result in a more refined list. In other words, the same dataset may yield different word lists depending on the filtering criteria used (Laosrirattanachai & Ruangjaroon, 2021a). Another limitation concerns the lexical units used in constructing the word list. While some studies employ word types (e.g., Bancroft-Billings, 2020) or word families (e.g., Greene & Coxhead, 2015), the present study uses lemmas as lexical units (e.g., *ANNOTATE*, which includes all inflected forms such as *annotate, annotates, annotated, and annotating*). Future research could explore the development of word lists based on word families. When word families are used as lexical units, all inflected and derived forms are included (e.g., *comprehend, comprehends, comprehended, comprehending, comprehension, comprehensions, and comprehensible*). However, when vocabulary is presented as word families, it may be more appropriate for upper-intermediate or advanced learners, as

learners at these proficiency levels are more likely to recognize the relationship between the headword (e.g., *comprehend*), its inflected forms within the same lemma, and its derived forms. As the third limitation, some technical vocabulary in LE may also appear in the GSL and AWL and was therefore excluded during lexical profiling. Future research could incorporate expert judgment to identify and reintroduce such technical terms in the LERC that were inadvertently removed because they overlap with the GSL and AWL. The last limitation of this corpus-based study concerns the genre in which LER vocabulary appears. In the current study, the data were drawn solely from journal articles, i.e., academic papers that represent research in LE. However, it should be noted that LER technical words can also be extracted from other sources, such as articles published in edited volumes, book chapters, as well as LE textbooks.

ACKNOWLEDGEMENTS

This study was supported by Thammasat University Research Fund, Contract No. TUFT 008/2568.

THE AUTHOR

Supakorn Phoocharoensil, PhD, is an Associate Professor of English at the Language Institute of Thammasat University (LITU), Thailand. He currently serves as the Director of LITU, and as Editor-in-Chief of *LEARN Journal: Language Education and Acquisition Research Network*. His research specializations include corpus linguistics for English language studies and ELT, corpus-based studies of synonymy, and second language acquisition.
supakorn.p@litu.tu.ac.th

REFERENCES

- Anthony, L. (2023). *AntWordProfiler*. [Software]. <http://www.laurenceanthony.net/software>
- Arunvong Na Ayutthaya, J., Kunthonjinda, K., Somwang, K., & Laosirattachai, P. (2022). Making beverage service word list for English for specific purposes classroom. *rEFLections*, 29(2), 325–343.
- Bancroft-Billings, S. (2020). Identifying spoken technical legal vocabulary in a law school classroom. *English for Specific Purposes*, 60, 9–25. <https://doi.org/10.1016/j.esp.2020.05.002>
- Brezina, V., & Gablasova, D. (2015). Is there a core general vocabulary? Introducing the new general service list. *Applied Linguistics*, 36(1), 1–22. <https://doi.org/10.1093/applin/amt018>
- Chung, T. M. (2003). A corpus comparison approach for terminology extraction. *Terminology*, 9(2), 221–245.
- Chung, T., & Nation, P. (2003). Technical vocabulary in specialized texts. *Reading in a Foreign Language*, 15(2), 103–116.
- Chung, T., & Nation, P. (2004). Identifying technical vocabulary. *System*, 32(2), 251–263. <https://doi.org/10.1016/j.system.2003.11.008>
- Cobb, T. (2021). *Web VocabProfile*. [Online program]. <http://www.lexutor.ca/vp/>
- Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, 34(2), 213–238. <https://doi.org/10.2307/3587951>

- Coxhead, A., & Hirsh, D. (2007). A pilot science-specific word list. *Revue Francaise de Linguistique Appliquee*, 12(2), 65–78. <https://doi.org/10.3917/rfla.122.0065>
- Coyle, D., Hood, P., & Marsh, D. (2010). The CLIL tool kit: Transforming theory into practice. In D. Coyle, P. Hood, & D. Marsh (Eds.), *CLIL: Content and language integrated learning* (pp. 45–73). Cambridge University Press.
- Dang, T. N. Y., & Webb, S. (2025). Applications of word lists in second language learning and teaching. *Language Teaching*, 58, 291–311. <https://doi.org/10.1017/S0261444825000059>
- Drašler, A., & Kavalir, M. (2026). Compiling a custom corpus and word list for ESAP: The case of English for Geographers. *English for Specific Purposes*, 81, 92–102. <https://doi.org/10.1016/j.esp.2025.09.004>
- Đurović, Z. (2021). Corpus linguistics methods for building ESP word lists, glossaries and dictionaries on the example of a Marine Engineering Word List. *Lexikos*, 31, 259–282.
- Flowerdew, L. (2012). *Corpora and language education*. Palgrave Macmillan.
- Gardner, D., & Davies, M. (2014). A new academic vocabulary list. *Applied Linguistics*, 35(3), 305–327. <https://doi.org/10.1093/applin/amt015>
- Greene, J. W., & Coxhead, A. (2015). *Academic vocabulary for middle school students: Research-based lists and strategies for key content areas*. Brookes Publishing.
- Hall, G. (2016). *The Routledge handbook of English language teaching*. Routledge.
- Hsu, W. (2014). Measuring the vocabulary load of engineering textbooks for EFL undergraduates. *English for Specific Purposes*, 33, 54–65. <https://doi.org/10.1016/j.esp.2013.07.001>
- Hyland, K., & Tse, P. (2007). Is there an “academic vocabulary”? *TESOL Quarterly*, 41(2), 235–253. <https://doi.org/10.1002/j.1545-7249.2007.tb00058.x>
- It-ngam, T., & Phoocharoensil, S. (2019). The development of science academic word list. *Indonesian Journal of Applied Linguistics*, 8(3), 657–667. <https://doi.org/10.17509/ijal.v8i3.15269>
- Khani, R., & Tazik, K. (2013). Towards the development of an academic word list for applied linguistics research articles. *RELC Journal*, 44(2), 209–232. <https://doi.org/10.1177/0033688213488432>
- Konstantakis, N. (2007). Creating a business word list for teaching business English. *Elia*, 7, 79–102.
- Kruawong, T., & Phoocharoensil, S. (2020). Developing an English zoology academic word list: A corpus-based study. *Thoughts*, 2020(2), 63–77.
- Kruawong, T., & Phoocharoensil, S. (2024). Enhancing science vocabulary and content knowledge of Thai EFL students through Content and Language Integrated Learning (CLIL) and English Science Textbooks Word Lists (ESTWL). *rEFlections*, 31(1), 275–294. <https://doi.org/10.61508/refl.v31i1.272797>
- Kruawong, T., & Phoocharoensil, S. (2025). Establishing English Science Textbook Word Lists (ESTWLs) for Thai EFL secondary students: A corpus-based approach. *PASAA*, 71(2), 386–418.
- Laosrirattanachai, P., & Laosrirattanachai, P. (2021). Applying lexical profiling to construct technical word lists for Thai tourist guides. *PASAA: A Journal of Language Teaching and Learning*, 62(1), 61–91.
- Laosrirattanachai, P., & Ruangjaroon, S. (2020). The word lists of hospitality service review construction. *Journal of Studies in the English Language*, 15(1), 107–158.
- Laosrirattanachai, P., & Ruangjaroon, S. (2021a). Corpus-based creation of tourism, hotel, and airline business word lists. *LEARN Journal: Language Education and Acquisition Research Network*, 14(1), 50–86.
- Laosrirattanachai, P., & Ruangjaroon, S. (2021b). Implementation of a data-driven hospitality lexis learning programme. *3L: The Southeast Asian Journal of English Language Studies*, 27(1), 1–21. <https://doi.org/10.17576/3L-2021-2701-01>
- Liu, J., & Han, L. (2015). A corpus-based environmental academic word list building and its validity test. *English for Specific Purposes*, 39, 1–11. <https://doi.org/10.1016/j.esp.2015.03.003>
- Mackey, A., & Gass, S. M. (2023). *Current approaches in second language acquisition research: A practical guide*. Wiley.

- Martinez, I., Beck, S., & Panza, C. (2009). Academic vocabulary in agriculture research articles: A corpus-based study. *English for Specific Purposes*, 28(3), 183–198. <https://doi.org/10.1016/j.esp.2009.04.001>
- Muñoz, V. (2015). The vocabulary of agriculture semi-popularization articles in English: A corpus-based study. *English for Specific Purposes*, 39, 26–44. <https://doi.org/10.1016/j.esp.2015.03.002>
- Namwong, B., Vorakhan, P., Suntorn, W., & Laosrirattanachai, P. (2022). Building a semi-technical political news word list for political science students. *Thoughts*, 2, 1–21.
- Nation, I. S. P. (2005). Teaching and learning vocabulary. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning*. Lawrence Erlbaum.
- Nation, I. S. P. (2006). How large a vocabulary is needed for reading and listening? *Canadian Modern Language Review*, 63(1), 59–82. <https://doi.org/10.3138/cmlr.63.1.59>
- Nation, I. S. P. (2016). *Making and using word lists for language learning and testing*. John Benjamins.
- Nation, I. S. P. (2022). *Learning vocabulary in another language* (3rd ed.). Cambridge University Press.
- Nation, I. S. P. & Webb, S. (2011). *Researching and analyzing vocabulary*. Heinle Cengage Learning.
- Nevisi, R. B., Safari, M., Hosseinpour, R. M., & Mousakazemi, R. S. (2023). A high frequency word list for political sciences. *Journal of Modern Research in English Language Studies*, 10(4), 21–43. <https://doi.org/10.30479/jmrels.2023.18213.2161>
- O’Keeffe, A., McCarthy, M., & Carter, R. (2007). *From corpus to classroom: Language use and language teaching*. Cambridge University Press.
- Ogden, C. K. (1930). *Basic English: A general introduction with rules and grammar*. Kegan Paul, Trench, Trubner & Co.
- Özer, M., & Akbaş, E. (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>
- Richards, I. A. (1943). *Basic English and its uses*. Kegan Paul, Trench, Trubner & Co.
- Rungrueang, T., Boonprasert, P., Poemongsajaroen, S., & Laosrirattanachai, P. (2022). Corpus-based approach to generate a word list for food service. *THAITESOL Journal*, 35(1), 57–76.
- Saed, H., Hussein, R. F., Haider, A. S., Al-Salman, S., & Oded, I. M. (2022). Establishing a COVID-19 lemmatized word list for journalists and ESP learners. *Indonesian Journal of Applied Linguistics*, 11(3), 577–588. <https://doi.org/10.17509/ijal.v11i3.37103>
- Schmitt, N., & Schmitt, D. (2012). How much vocabulary is needed to use English? Replication of Nation (2006) and beyond. *Language Teaching*, 45(3), 319–332.
- Szudarski, P. (2018). *Corpus linguistics for vocabulary: A guide for research*. Routledge.
- Tangpijaikul, M. (2014). Preparing business vocabulary for the ESP classroom. *RELJ Journal*, 45(1), 51–65. <https://doi.org/10.1177/0033688214522641>
- Valipouri, L., & Nassaji, H. (2013). A corpus-based study of academic vocabulary in chemistry research articles. *English for Specific Purposes*, 12(4), 248–263. <https://doi.org/10.1016/j.esp.2013.07.001>
- Van Zeeland, H., & Schmitt, N. (2013). Lexical coverage in L1 and L2 listening comprehension: The same or different from reading comprehension? *Applied Linguistics*, 34(4), 457–479. <https://doi.org/10.1093/applin/ams074>
- Ward, N. (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.2009.04.003>
- Watson Todd, R. (2017). An opaque engineering word list: Which words should a teacher focus on? *English for Specific Purposes*, 45(1), 31–39. <https://doi.org/10.1016/j.esp.2016.09.001>
- Webb, S., & Nation, P. (2017). *How vocabulary is learned*. Oxford University Press.
- Wei, L., Hua, Z., & Simpson, J. (2024). *The Routledge handbook of applied linguistics*. Routledge.



- West, M. (1953). *A general service list of English words*. Longman.
- 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), Article 32. <https://doi.org/10.1186/s40862-023-00205-5>
- Yang, M. (2015). A nursing academic word list. *English for Specific Purposes*, 37(1), 27–38. <https://doi.org/10.1016/j.esp.2014.05.003>

Appendix

The Language Education Research Word List (LERWL)

No.	Words				
1	A2	51	cognition	102	Efficacy
2	accent	52	cognitive	103	EFL
3	adjective	53	cohesion	104	EL
4	administer	54	collaboration	105	Electronic
5	affordance	55	collaborative	106	Elementary
6	AI	56	collaboratively	107	ELF
7	AL	57	collocate	108	elicit
8	Align	58	collocation	109	ELP
9	Alignment	59	competence	110	ELS
10	American	60	competency	111	ELT
11	Annotate	61	competent	112	embed
12	annotation	62	comprehensible	113	EMI
13	ANOVA	63	comprehension	114	emotion
14	aptitude	64	comprehend	115	emotional
15	Arabic	65	conception	116	engage
16	argumentation	66	conceptualizationn	117	engagement
17	argumentative	67	concrete	118	Englishes
18	Asian	68	conjunction	119	entail
19	asynchronous	69	consonant	120	epistemic
20	audio	70	continuum	121	ER
21	auditory	71	conversational	122	ES
22	aural	72	corpus	123	ESL
23	authentic	73	corrective	124	ESP
24	authenticity	74	correlation	125	essay
25	autonomous	75	CPD	126	exam
26	autonomy	76	CR	127	excerpt
27	B1	77	credibility	128	facial
28	B2	78	Cronbach	129	feedback
29	baseline	79	cue	130	fixation
30	bilingual	80	cumulative	131	flip
31	bilingualism	81	curriculum	132	fluency
32	binary	82	DA	133	fluent
33	blend	83	dataset	134	formative
34	British	84	deductive	135	formulaic
35	burnout	85	demotivation	136	foster
36	CAF	86	descriptor	137	frustration
37	C1	87	developmental	138	genre
38	C2	88	diagnostic	139	gloss
39	causative	89	dialogic	140	grammatical
40	CEFR	90	dialogue	141	graphic
41	chat	91	dialog	142	grit
42	Chinese	92	discourse	143	guidance
43	classifier	93	discursive	144	holistic
44	classroom	94	digital	145	homework
45	CLIL	95	dual	146	homogeneity
46	CLT	96	Dutch	147	iBT
47	cluster	97	dyad	148	ICC
48	CMC	98	EAP	149	idiom
49	coder	99	ecological	150	IELTS
50	coefficient	100	ED	151	immersion
		101	EFA	152	incidental

153	index
154	Indonesian
155	inductive
156	inferential
157	instructional
158	intelligibility
159	interactional
160	intercultural
161	interlocutor
162	interpersonal
163	interplay
164	interpreter
165	interview
166	intonation
167	inventory
168	Iranian
169	Italian
170	Japanese
171	Korean
172	L1
173	L2
174	LANG
175	latent
176	lexical
177	lexis
178	Likert
179	linguistic
180	linguistics
181	literacy
182	longitudinal
183	Mandarin
184	mainstream
185	major
186	map
187	metalinguistic
188	metalinguistics
189	metaphor
190	metaphorical
191	metric
192	micro
193	mobile
194	modal
195	modality
196	monolingual
197	morphological
198	motivational
199	multilingual
200	multimedia
201	multimodal
202	narrative
203	negation
204	negotiate
205	neural
206	NNS

207	novice
208	NS
209	objective
210	online
211	optimal
212	oral
213	orientation
214	orthographic
215	outperform
216	paraphrase
217	Pearson
218	pedagogical
219	pedagogy
220	peer
221	Persian
222	personality
223	phonetic
224	phonetics
225	phonological
226	phrase
227	pilot
228	placement
229	platform
230	portfolio
231	Portuguese
232	posttest
233	practice
234	practise
235	practicum
236	pragmatic
237	pragmatics
238	predictive
239	predictor
240	presentation
241	pretest
242	proficiency
243	proficient
244	pronoun
245	pronunciation
246	punctuation
247	quantitative
248	quasi
249	query
250	questionnaire
251	quiz
252	rater
253	rating
254	rationale
255	readability
256	recall
257	recast
258	receptive
259	recording
260	reference

261	regression
262	repertoire
263	repetition
264	resilience
265	resiliency
266	rhetorical
267	rubric
268	saliency
269	saliency
270	scaffold
271	score
272	script
273	segment
274	semantic
275	semantics
276	semester
277	semiotic
278	sentiment
279	setting
280	sig
281	simultaneously
282	singular
283	SLA
284	sociocultural
285	software
286	Spanish
287	spontaneous
288	SPSS
289	stakeholder
290	stance
291	standardize
292	stimulate
293	stimulus
294	subjective
295	summative
296	supervise
297	Swedish
298	switch
299	syllable
300	syllabus
301	synchronous
302	syntactic
303	syntax
304	synthesis
305	taxonomy
306	TBLT
307	telecollaboration
308	temporal
309	tertiary
310	TESOL
311	textbook
312	threshold
313	TOEFL
314	TOEIC

315	token
316	trait
317	trajectory
318	transcribe
319	transcription
320	translanguaging
321	translator
322	TTS
323	Turkish
324	tutor

325	tweet
326	twitter
327	understanding
328	usage
329	utterance
330	variance
331	verbal
332	video
333	Vietnamese
334	VLS

335	vocabulary
336	washback
337	WCF
338	whatsapp
339	workshop
340	WTC
341	youtube
342	ZPD