



Need Analysis of A Corpus-Based Electronic Games To Improve Vocabulary Of Junior High School Students

Asti Ramadhani Endah Lestari

Universitas Negeri Jakarta, Indonesia

Corresponding Author: asti.ramadhani@unj.ac.id

Abstract. Need analysis plays a vital role in designing a learning media as it helps to decide a suitable contents and features with the target users. The present study describes the needs of Junior High School students and teachers for a corpus-based electronic vocabulary games. A questionnaire adapting Hutchinson and Waters model of Need Analysis was administered to 111 students. The questionnaire investigated students' needs as regards their attitude, learning experience, preferred content, and preferred feature for an android-based vocabulary game. Then, the survey was followed by in-depth interview with students and teachers. The findings showed that the teacher tried to use online games in the classroom and it increased students' motivation. However, Teachers found out that existing games could not accommodate the learning process because the vocabulary provided was not suitable with students' ability. In addition, students like to play games that can be accessed through android-based mobile phones and consists of challenges with different missions in each level. The result of the study is beneficial as the basic consideration in further research to develop a corpus-based electronic games.

Keywords: Corpus; Gamifying; Need Analysis; TEFL; Vocabulary.

1. INTRODUCTION

Vocabulary knowledge is widely recognized as a fundamental component of foreign language learning because it serves as the foundation for effective communication. Without sufficient vocabulary mastery, learners face difficulties in understanding and expressing ideas, regardless of their grammatical competence. Vocabulary learning entails not only understanding the meaning of words but also acquiring knowledge of pronunciation, spelling, collocations, grammatical behavior, and contextual usage (Nation, 2001). Vocabulary size and depth have been consistently associated with learners' overall language proficiency, making vocabulary development a crucial aspect of English language learning (Nation & Waring, 1997). Research has also revealed that learners often perceive vocabulary as one of the most important elements of language acquisition and express a need for more intensive vocabulary practice to enhance their language skills.

To facilitate vocabulary development, teachers are expected to employ effective instructional strategies and select vocabulary items that are most relevant to learners' communicative needs. Previous studies have suggested that vocabulary learning becomes more meaningful when learners encounter words in authentic contexts and use them in communicative situations (Nation, 2001). In recent years, technology-enhanced language learning has attracted considerable attention as an alternative approach to vocabulary instruction. Studies have demonstrated that technology-mediated learning environments can improve vocabulary acquisition and learner engagement compared to traditional instructional approaches. Among these innovations, gamification has emerged as a promising strategy for

promoting motivation and participation in language learning. Gamification refers to the use of game design elements in non-game contexts (Deterding et al., 2011). In educational settings, gamification incorporates elements such as points, badges, levels, challenges, rewards, and immediate feedback to increase learner engagement and motivation (Christopoulos & Mystakidis, 2023). Previous studies have reported that gamified learning environments can reduce learners' anxiety, increase learning enjoyment, and improve vocabulary retention. These benefits are particularly important in vocabulary learning because learners often perceive vocabulary memorization as a demanding and repetitive activity. Therefore, gamified learning environments have the potential to foster more positive attitudes toward vocabulary acquisition.

Despite the potential advantages of various instructional approaches, the effectiveness of vocabulary learning largely depends on the relevance of the vocabulary being taught. Not all vocabulary items contribute equally to learners' communicative goals. Consequently, teachers need to prioritize vocabulary that is useful, frequently encountered, and aligned with learners' educational and professional needs. Nation (2001) argued that language learners benefit more from mastering a relatively small number of high-frequency words than from studying large quantities of low-frequency vocabulary. Similarly, Nation and Waring (1997) emphasized that high-frequency vocabulary provides the greatest coverage of written and spoken texts, making it essential for language learners.

Corpus linguistics offers a systematic approach to identifying useful vocabulary for language instruction. A corpus consists of a large collection of authentic spoken and written texts that represent actual language use in real communicative contexts (McEnery & Hardie, 2012). Through corpus analysis, researchers can identify patterns of language use, word frequencies, collocations, and discourse characteristics. Because corpora are derived from authentic communication, they provide reliable empirical evidence for vocabulary selection and syllabus design. Furthermore, corpus-based learning supports authenticity, learner autonomy, and data-driven learning by allowing learners to discover language patterns independently from real-world language data (Johns, 1991).

Corpus research has also contributed significantly to the development of frequency-based vocabulary lists. One of the earliest and most influential vocabulary lists was the *General Service List* (GSL) developed by West (1953), which contains approximately 2,000 high-frequency English words. The GSL was designed to identify vocabulary items of greatest value to English language learners and has served as the foundation for numerous subsequent vocabulary lists. Research indicates that mastery of these high-frequency words provides

substantial coverage of everyday spoken and written English (West, 1953; Nation, 2013). Building upon the GSL, Nation developed frequency-based vocabulary classifications that have become widely used in language teaching, curriculum design, and vocabulary assessment.

While corpus-based vocabulary lists provide valuable information about word frequency and usefulness, effective vocabulary instruction should also consider learners' actual needs. Needs analysis is regarded as a fundamental stage in curriculum and course design because it enables educators to identify learners' goals, preferences, proficiency levels, and learning challenges (Hutchinson & Waters, 1987; Dudley-Evans & St John, 1998). Through needs analysis, teachers can determine what learners need to learn, what they already know, and what instructional approaches are most suitable for them. As a result, learning materials and activities can be designed to better align with learners' expectations and educational objectives.

Although previous studies have examined the benefits of gamification for vocabulary learning and the application of corpus linguistics in vocabulary selection, relatively few studies have investigated learners' vocabulary needs as a basis for developing corpus-based and technology-enhanced vocabulary instruction. Understanding students' needs is particularly important because the effectiveness of vocabulary learning resources depends not only on word frequency but also on learners' academic requirements, learning preferences, and perceived difficulties. Therefore, a comprehensive needs analysis is necessary before designing vocabulary learning materials and activities.

Accordingly, this study aims to analyze students' vocabulary learning needs in order to provide a foundation for the development of corpus-based vocabulary learning materials and gamified instructional activities. Specifically, the study investigates students' current vocabulary learning practices, perceived difficulties, target vocabulary needs, preferred learning methods, and expectations regarding technology-assisted vocabulary learning. The findings are expected to contribute to the development of learner-centered vocabulary instruction that combines corpus-based evidence with students' actual learning needs.

2. LITERATURE REVIEW

Vocabulary Mastery

Vocabulary mastery is a fundamental component of language proficiency and one of the most important learning objectives for foreign language learners. A sufficient vocabulary repertoire enables learners to comprehend spoken and written texts, communicate effectively, and participate successfully in academic activities. According to Yang and Coxhead (2020),

vocabulary development is influenced by several key factors, including word frequency, vocabulary load, and repetition. These factors determine the extent to which learners encounter and retain lexical items in different learning contexts. Furthermore, Kamroto et al. (2022) emphasized that vocabulary knowledge should be examined in relation to text coverage, which refers to the proportion of words in a text that are known by learners. Adequate lexical coverage is essential for successful comprehension because learners may experience difficulty understanding texts when too many unfamiliar words are encountered.

Word frequency is one of the most significant considerations in vocabulary learning and instruction. Frequency refers to how often a word occurs in a language corpus and is generally categorized into high-frequency, mid-frequency, and low-frequency vocabulary (Love, 2020). High-frequency vocabulary consists of the first and second 1,000 word families and includes most function words and many common content words. These words account for the largest proportion of language use in both spoken and written texts and are therefore considered essential for language learners (Nation, 2013). In contrast, mid-frequency vocabulary encompasses approximately the third to ninth 1,000 word families. Knowledge of these words enables learners to comprehend a wider range of authentic texts and contributes significantly to reading fluency and language development. Low-frequency vocabulary consists of words beyond the ninth 1,000 word families and occurs relatively infrequently in language use. Although these words represent a large number of lexical items, they contribute only a small proportion of the total vocabulary found in most texts (Nation, 2013).

Another important aspect of vocabulary mastery is vocabulary load, which refers to the lexical demands imposed by a text. The concept of vocabulary load is closely related to lexical coverage, defined as the percentage of running words (tokens) in a text that are already known by learners (Yang & Coxhead, 2020). Research suggests that learners require approximately 95% lexical coverage for adequate comprehension and up to 98% coverage for optimal comprehension of authentic texts (Nation, 2006). Therefore, understanding vocabulary load is crucial for selecting instructional materials that correspond to learners' proficiency levels and support effective vocabulary acquisition.

Overall, vocabulary mastery involves more than simply memorizing word meanings. It encompasses knowledge of word frequency, contextual usage, collocations, and the ability to comprehend and produce vocabulary appropriately in different communicative situations. Consequently, vocabulary instruction should prioritize high-frequency and pedagogically useful words that can maximize learners' comprehension and language use.

Vocabulary Games

The integration of games into language learning has gained considerable attention due to its potential to enhance learner engagement and motivation. In vocabulary instruction, games are particularly valuable because they create an enjoyable learning environment while simultaneously supporting vocabulary retention and recall. Hasram et al. (2021) argued that vocabulary games strengthen learners' memory by providing repeated exposure to lexical items in meaningful and interactive contexts. Through game-based activities, learners are encouraged to actively process vocabulary rather than passively memorize word lists, which can improve long-term retention.

Recent developments in educational technology have expanded the use of digital and online games in vocabulary learning. Chapman and Rich (2018) defined game-based learning as the integration of game thinking and game mechanics into instructional activities to enhance learner participation and achievement. Common game elements include points, levels, challenges, leaderboards, rewards, and immediate feedback. These elements create a sense of accomplishment and competition that can increase learners' intrinsic motivation to engage in vocabulary learning activities.

The rapid growth of mobile technology has further facilitated the use of game-based language learning. Daud et al. (2015) reported that mobile-assisted language learning environments enable learners to access vocabulary activities anytime and anywhere, thereby increasing opportunities for language exposure and practice. Mobile learning applications also support personalized learning experiences through adaptive exercises and instant feedback.

Several empirical studies have demonstrated the effectiveness of vocabulary games in improving language learning outcomes. Andreani and Ying (2019), for example, found that interactive online games significantly enhanced students' vocabulary learning experiences, particularly among learners with lower levels of English proficiency. The interactive nature of digital games promotes active participation, reduces learning anxiety, and encourages learners to practice vocabulary repeatedly in meaningful contexts. These findings suggest that game-based vocabulary learning can serve as an effective instructional strategy for supporting vocabulary development among English as a Foreign Language (EFL) learners.

Corpus-Based Vocabulary Learning

Corpus linguistics has become an increasingly important field in language education due to its ability to provide authentic evidence of language use. A corpus is a large, systematic collection of spoken or written texts compiled for linguistic analysis. According to Rogers et al. (2021), corpus data serve as a powerful resource for developing language learning materials

based on real-world language usage rather than intuition or prescriptive descriptions. Because corpora reflect authentic communication, they provide valuable information about word frequency, collocations, phraseology, and contextual patterns. One of the primary advantages of corpus-based vocabulary learning is that it allows educators to identify lexical items that learners are most likely to encounter in actual language use. Frequency information derived from corpora helps teachers prioritize vocabulary that has the greatest communicative value. High-frequency vocabulary lists developed from corpus data have therefore become important resources for curriculum design, vocabulary assessment, and instructional material development (Nation, 2013).

Corpus-based approaches also promote data-driven learning, in which learners explore authentic language examples and discover linguistic patterns independently. Through exposure to concordance lines and real contexts of use, learners develop a deeper understanding of vocabulary meaning, collocation, and grammatical behavior. As a result, corpus-based learning encourages learner autonomy and supports a more inductive approach to language acquisition.

Previous studies have predominantly applied corpus-based vocabulary development in higher education contexts. For instance, Xodabande et al. (2020) demonstrated the effectiveness of corpus-informed vocabulary instruction in improving university students' lexical knowledge and academic language proficiency. Similarly, Kilgariff et al. (2014) highlighted the pedagogical value of corpus tools in creating authentic learning materials and facilitating vocabulary analysis for language learners. Despite the growing body of research on corpus-based learning, its application in vocabulary instruction remains predominantly focused on academic contexts and tertiary education settings. Therefore, further research is needed to explore how corpus-based vocabulary resources can be adapted to meet learners' specific needs and integrated with innovative instructional approaches such as game-based learning.

3. RESEARCH METHOD

Research Design and Participants

This study employed the needs analysis phase of a Research and Development (R&D) design to investigate students' needs as a preliminary step in developing a corpus-based electronic game for English vocabulary learning. The needs analysis aimed to identify learners' perceptions of English learning, vocabulary learning difficulties, technology usage, gaming experiences, and expectations regarding the proposed learning model. The results of this phase were used to establish the pedagogical and technological specifications of the product to be developed.

The participants were 111 junior high school students enrolled in schools in East Jakarta, Indonesia. The participants represented the target users of the proposed corpus-based electronic vocabulary game. They were selected based on accessibility and participation in the online needs analysis survey. In addition, several English teachers and student representatives participated in Focus Group Discussions (FGDs) to provide deeper insights into vocabulary learning needs and instructional practices.

Data Collection

Data were collected using three instruments: a questionnaire, Focus Group Discussions (FGDs), and document analysis. The primary instrument was a needs analysis questionnaire administered online through Google Classroom. The questionnaire consisted of closed-ended and open-ended items designed to investigate students' perceptions of English learning, vocabulary learning challenges, technology access, experience with electronic games, and expectations for an electronic game-based learning application. The closed-ended items utilized a five-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree).

To complement the questionnaire findings, Focus Group Discussions (FGDs) were conducted with selected students and English teachers. The discussions explored participants' experiences in vocabulary learning, difficulties encountered during English classes, preferred learning activities, and expectations regarding the implementation of a corpus-based electronic game. A third instrument involved document analysis, which examined curriculum documents, English textbooks, learning objectives, competency standards, and instructional materials used at the junior high school level. This analysis was conducted to identify vocabulary items that should be included in the limited corpus used as the linguistic foundation of the proposed game. Data collection was conducted in four stages. First, the questionnaire was distributed through Google Classroom, allowing students to complete the survey electronically. The responses were automatically recorded and compiled for analysis.

Second, Focus Group Discussions were organized with English teachers and selected students. These discussions provided qualitative data regarding learners' needs, vocabulary learning difficulties, and preferences for digital learning media.

Third, curriculum and instructional documents were reviewed to identify vocabulary learning objectives and high-priority lexical items relevant to junior high school students. The findings from this stage were used to establish a limited corpus consisting of vocabulary that aligned with curricular requirements and students' learning needs. Finally, the researchers conducted an analysis of existing electronic vocabulary-learning games. The analysis focused on instructional content, game mechanics, feedback systems, reward structures, user interface

design, and accessibility features. The results were used to identify effective design elements that could be incorporated into the development of the corpus-based electronic game. The quantitative data obtained from the questionnaire were analyzed using descriptive statistics, including frequencies, percentages, means, and standard deviations. Meanwhile, qualitative data from the FGDs, open-ended responses, and document analysis were analyzed thematically to identify recurring patterns related to learners' vocabulary needs, learning preferences, and expectations for technology-assisted vocabulary learning.

4. RESULTS AND DISCUSSION

Students' Perceptions of English Learning

The first part of the questionnaire explored students' perceptions of English learning. The findings indicated that most respondents viewed English as an important subject that could support their academic achievement and future opportunities. Students also acknowledged the importance of vocabulary mastery as a foundation for developing other language skills such as reading, writing, listening, and speaking.

However, despite their positive attitudes toward English, many students reported experiencing difficulties in language learning. Vocabulary emerged as one of the most challenging aspects of English learning. Students frequently reported problems understanding unfamiliar words, remembering newly learned vocabulary, and using vocabulary appropriately in context. This finding supports Nation's (2013) assertion that vocabulary knowledge is central to language proficiency and often becomes a major obstacle for language learners who possess limited lexical resources.

Table 1. Students' Perceptions of English Learning (N = 111).

Statement	Disagree (1-2)	Neutral (3)	Agree (4-5)
English is beneficial for daily life	0 (0%)	6 (5.4%)	105 (94.6%)
English helps me learn other subjects	2 (1.8%)	10 (9.0%)	99 (89.2%)
English is a subject that I like	22 (19.8%)	39 (35.1%)	50 (45.0%)
Vocabulary is the most difficult aspect of English learning	29 (26.1%)	46 (41.4%)	36 (32.4%)

The findings indicate that students generally hold positive attitudes toward English learning. Approximately 94.6% of respondents agreed that English is useful in their daily lives, while 89.2% believed that English contributes to learning other school subjects. These findings suggest that students recognize the importance of English as an academic and practical skill. Nevertheless, vocabulary remains a significant challenge, with almost one-third of respondents identifying it as one of the most difficult components of English learning. This finding supports

the view that vocabulary knowledge is central to language proficiency and requires targeted instructional support.

Table 2. Difficulties in English Learning (N = 111).

Skill Aspect	Agree (4-5)	Neutral (3)	Disagree (1-2)
Listening is difficult	34 (30.6%)	40 (36.0%)	37 (33.3%)
Speaking is difficult	50 (45.0%)	35 (31.5%)	26 (23.4%)
Reading is difficult	20 (18.0%)	41 (36.9%)	50 (45.0%)
Writing is difficult	12 (10.8%)	35 (31.5%)	64 (57.6%)

Speaking (45.0%) and grammar (40.5%) were identified as the most challenging aspects of English learning, followed by vocabulary (32.4%). Although vocabulary was not ranked as the most difficult skill area, it was consistently identified as a learning obstacle. Since vocabulary supports all language skills, limitations in vocabulary knowledge may contribute to difficulties in speaking, listening, reading, and writing.

Table 3. Students' Attitudes toward Learning Resources and Technology (N = 111).

Statement	Disagree n (%)	Neutral n (%)	Agree n (%)
Learning English only from textbooks is sufficient	74 (66.7)	22 (19.8)	15 (13.5)
I can learn English from resources other than textbooks	16 (14.4)	25 (22.5)	70 (63.1)
I enjoy reading English books, websites, or social media	28 (25.2)	36 (32.4)	47 (42.3)
I enjoy listening to English songs	8 (7.2)	13 (11.7)	90 (81.1)
I enjoy playing electronic games in English	18 (16.2)	23 (20.7)	70 (63.1)
Learning through games is more interesting than traditional activities	14 (12.6)	21 (18.9)	76 (68.5)
English learning applications can help me improve vocabulary	10 (9.0)	23 (20.7)	78 (70.3)

Two-thirds of the participants believed that textbooks alone were insufficient for learning English. Additionally, 63.1% reported enjoying English electronic games, while 81.1% enjoyed listening to English songs. These findings indicate that students prefer learning experiences beyond traditional classroom materials. Therefore, integrating technology and game-based learning elements could increase learners' engagement and motivation.

The results revealed that most students perceived English as an important subject for both academic and practical purposes. More than 90% of respondents agreed that English is beneficial for daily life and helps them learn other subjects. These findings suggest that students are aware of the value of English as an international language and recognize its role in educational success. However, despite these positive attitudes, students reported experiencing considerable difficulties in English learning, particularly in vocabulary, grammar, and speaking. Vocabulary was identified as a significant learning challenge by approximately one-

third of the respondents. This finding supports Nation's (2013) argument that vocabulary knowledge constitutes the foundation of language competence because learners cannot effectively communicate, read, write, or understand spoken language without sufficient lexical resources. The findings further support Yang and Coxhead's (2020) proposition that vocabulary acquisition depends on repeated exposure to vocabulary items and opportunities for contextualized practice. Students' difficulties in remembering and using vocabulary indicate that the existing learning environment may not provide sufficient opportunities for meaningful encounters with lexical items. Therefore, vocabulary instruction should move beyond rote memorization toward activities that facilitate repeated exposure and contextualized language use.

The results also align with Kamroto et al. (2022), who emphasized the importance of vocabulary coverage for language comprehension. Students' reported difficulties in understanding texts and classroom materials may be attributed to insufficient vocabulary knowledge, which limits their ability to achieve adequate lexical coverage during reading and listening activities. One notable finding was that 66.7% of students believed that learning English through textbooks alone was insufficient. Furthermore, a considerable number of students reported that traditional vocabulary learning activities were monotonous and less motivating. These findings suggest a mismatch between students' learning preferences and current instructional practices.

Contemporary language learning theories emphasize the importance of learner engagement in promoting language acquisition. Vocabulary learning is more effective when learners actively interact with language rather than passively memorize isolated word lists (Nation, 2013). The students' responses indicate a need for learning environments that provide interaction, challenge, and meaningful language practice. The present findings suggest that students prefer multimodal learning experiences involving visual, auditory, and interactive components. This preference is evident from the high percentage of students who enjoy learning through songs, online resources, and digital applications. Such findings reinforce previous studies demonstrating that technology-enhanced language learning can increase learner engagement and create additional opportunities for vocabulary exposure outside the classroom.

Moreover, the findings highlight the importance of learner-centered instructional design. According to Hutchinson and Waters (1987), needs analysis serves as the foundation for developing instructional materials that align with learners' expectations and learning requirements. The students' expressed desire for more engaging learning media demonstrates

the necessity of incorporating their perspectives into the design of future vocabulary learning tools. An important finding of the study is that students prefer learning vocabulary that is useful for communication and relevant to their academic needs. This finding highlights the importance of selecting vocabulary systematically rather than relying solely on textbook content.

Corpus linguistics provides a strong theoretical framework for addressing this issue. According to Rogers et al. (2021), corpus-based approaches enable educators to identify vocabulary that learners are most likely to encounter in authentic language use. Through corpus analysis, vocabulary selection can be based on empirical evidence rather than intuition. The findings from the needs analysis support Nation's (2013) argument that vocabulary instruction should prioritize high-frequency vocabulary because such words provide the greatest coverage of authentic texts. Similarly, Love (2020) emphasized that high-frequency words form the core vocabulary needed for effective communication. Consequently, incorporating corpus-based high-frequency vocabulary into the proposed game would ensure that students focus on lexical items with the greatest learning value.

The use of a limited corpus derived from curriculum documents and frequency-based vocabulary lists offers additional advantages. First, it ensures alignment between learning materials and curriculum objectives. Second, it increases the likelihood that students will encounter the target vocabulary in future language-learning contexts. Finally, it supports the principles of data-driven learning by exposing students to authentic language patterns and contexts. This finding is particularly important because previous corpus-based vocabulary studies have focused primarily on higher education settings (Kilgarriff et al., 2014; Xodabande et al., 2020). The present study suggests that corpus-based approaches also have strong potential for supporting vocabulary learning among junior high school students.

5. CONCLUSIONS

Based on the needs analysis, the proposed learning model should incorporate four key components. First, vocabulary content should be selected from a limited corpus consisting of high-frequency and curriculum-relevant lexical items. Second, learning activities should emphasize contextualized vocabulary use rather than isolated memorization. Third, the application should integrate gamification elements such as points, badges, levels, challenges, rewards, and immediate feedback. Finally, the platform should be optimized for Android smartphones to maximize accessibility among target users. Overall, the findings suggest that a corpus-based electronic game has considerable potential to address students' vocabulary-learning needs while simultaneously increasing motivation, engagement, and opportunities for

autonomous learning. By combining principles from corpus linguistics, vocabulary acquisition theory, and gamification, the proposed learning model may provide a more effective alternative to traditional vocabulary instruction for junior high school learners.

REFERENCES

- Andreani, W., & Ying, Y. C. (2019). Powtoon as a tool to enhance students' learning experience and creativity in vocabulary learning. *Journal of English Language Teaching and Linguistics*, 4(2), 203–221.
- Chapman, J. R., & Rich, P. J. (2018). Does educational gamification improve students' motivation? If so, which game elements work best? *Journal of Educational Technology Systems*, 46(3), 315–331. <https://doi.org/10.1177/0047239517718518>
- Christopoulos, A., & Mystakidis, S. (2023). Gamification in education: A systematic literature review of empirical evidence. *Education Sciences*, 13(7), Article 760. <https://doi.org/10.3390/educsci13070760>
- Daud, N. M., Abdul Karim, N. S., Hassan, N., & Rahman, N. A. (2015). Mobile-assisted language learning in improving English vocabulary among ESL learners. *International Journal of Social Science and Humanity*, 5(6), 573–577.
- Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011). From game design elements to gamefulness: Defining gamification. In *Proceedings of the 15th International Academic MindTrek Conference* (pp. 9–15). Association for Computing Machinery. <https://doi.org/10.1145/2181037.2181040>
- Dudley-Evans, T., & St John, M. J. (1998). *Developments in English for specific purposes: A multi-disciplinary approach*. Cambridge University Press.
- Hasram, S., Rahman, A. A., & Yusof, N. (2021). The effectiveness of vocabulary games in enhancing students' vocabulary mastery. *Journal of Language and Linguistic Studies*, 17(2), 1101–1114.
- Hutchinson, T., & Waters, A. (1987). *English for specific purposes: A learning-centred approach*. Cambridge University Press.
- Johns, T. (1991). Should you be persuaded: Two examples of data-driven learning. In T. Johns & P. King (Eds.), *Classroom concordancing* (pp. 1–16). ELR Journal.
- Kamrotov, M., Bui, G., & Nation, I. S. P. (2022). Vocabulary knowledge and lexical coverage: Implications for second language reading comprehension. *Reading in a Foreign Language*, 34(1), 1–23.
- Kilgariff, A., Baisa, V., Bušta, J., Jakubíček, M., Kovář, V., Michelfeit, J., Rychlý, P., & Suchomel, V. (2014). The Sketch Engine: Ten years on. *Lexicography*, 1(1), 7–36. <https://doi.org/10.1007/s40607-014-0009-9>
- Love, R. (2020). *Corpus linguistics and language variation*. Routledge. <https://doi.org/10.4324/9780429020864>
- McEnery, T., & Hardie, A. (2012). *Corpus linguistics: Method, theory and practice*. Cambridge University Press.
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge University Press.

- 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. (2013). *Learning vocabulary in another language* (2nd ed.). Cambridge University Press.
- Nation, P., & Waring, R. (1997). Vocabulary size, text coverage and word lists. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy* (pp. 6–19). Cambridge University Press.
- Rogers, J., Webb, S., & Nakata, T. (2021). Corpus-informed vocabulary learning and teaching. *Language Teaching Research Quarterly*, 21, 1–16.
- West, M. (1953). *A general service list of English words*. Longman.
- Xodabande, I., Iravi, Y., & Shabani, K. (2020). The effectiveness of corpus-based vocabulary instruction on EFL learners' lexical development. *Computer Assisted Language Learning Electronic Journal*, 21(2), 191–211.
- Yang, Y., & Averil, C. (2020). Vocabulary learning, frequency, and lexical coverage.
- Yang, Y., & Coxhead, A. (2020). The role of vocabulary load and lexical coverage in second language reading comprehension. *TESOL Quarterly*, 54(4), 1068–1090.