

Corpus Linguistics and English Grammar Instruction in Thai Educational Context

Nuntiporn Raungsawat¹

Abstracts

Teaching grammar is becoming a challenge for a language teacher. Grammar is a foundational component of language instruction which should be well-planned and delivered to the learners. Therefore, the teacher should carefully select teaching approach to enhance students' grammatical competency. In Thai EFL context, English grammar has long been taught in a traditional way, where the teacher presents rules, followed by providing examples and translation exercises. Recently, corpus linguistics has been introduced as an alternative approach to teaching grammar, as it enables learners to discover and analyze complex patterns of language structures used in real-life situations. Through data-driven learning, grammar teaching can be updated, accurate, and aligned with the real-world usage of the English language. This article primarily aims to provide a comprehensive overview of corpus linguistics and its pedagogical implications in grammar instruction, particularly in Thai EFL classrooms. It begins with the importance of grammar in language learning and English grammar instruction. Then, the definitions of corpus linguistics and corpora are provided. Additionally, previous studies on the investigation of the corpus-based approach in grammar teaching conducted in Thai educational contexts are reviewed. Furthermore, the benefits and challenges of corpora in English grammar instruction are illustrated. The conceptual framework related to the topic is presented as the insights from the literature review. Finally, the paper suggests specific steps for implementing a corpus-based approach in English grammar pedagogy.

Keywords: Corpus Linguistics, English Grammar Instruction, Thai Educational Context

¹ Lecturer, Ph.D. in English Language Teaching, English Department, Faculty of Humanities and Social Sciences, Rajabhat Rajanagarindra University, Thailand
Corresponding Author E-mail: nuntiporn.rau@gmail.com

Introduction

English is an international language that people worldwide use to communicate with each other. According to Hymes (1972), language learners need communicative competence, including grammatical, sociolinguistic, discourse, and strategic competence, in order to communicate appropriately in different social contexts. Based on this notion, grammar knowledge is one of the crucial aspects of English language learning. Saengboon et al. (2022) describe grammar as “the backbone of a strong linguistic body” (p.187). Learners who possess a solid foundation in English grammar can enhance their academic achievement, communicative competence, and have better career opportunities after graduation (Sukavatee, 2023). Moreover, Phoocharoensil (2023) adds that learners should not ignore the importance of grammar because it is influential on success in second language communication. According to Azar (2007), grammar is “the weaving that creates the fabric” (p. 2). Without grammar, language learners would use only individual words or sounds, and body gestures to communicate meaning. Simply put, EFL students are required to understand the grammatical structure rules as well as apply those rules correctly for a variety of communicative purposes.

Despite several advantages of grammar, previous studies have indicated that insufficient knowledge of grammatical structures is one of the most common problems in learning English for Thai students at all levels. For example, Kaosayapandhu (2023) claimed that Thai students at the primary and secondary school levels were unable to communicate effectively in English due to inadequate knowledge of grammar. Yuh and Kaewurai (2021) found that Thai students in a secondary school reported inability to use correct grammatical structures as the most serious problem in their English speaking. Tantiwich and Sinwongsuwan (2021) revealed that Thai EFL university students lacked sufficient knowledge to construct grammatically correct English sentences in their conversations. Kampookaew (2020) discovered that Thai EFL university students typically made grammatical errors in an EAP writing task. One of the main reasons behind the cause of inadequate English grammatical knowledge of Thai students probably stems from the teacher-centered approach which focuses more heavily on grammatical rules but pays less attention to interactive activities for students to practice their communication skills (Yuh & Kaewurai, 2021). Sukavatee (2023) depicts this conventional teaching style as “a spoon-feeding approach” (p. 274). To clarify, the teacher provides all the knowledge to the class, and students only listen and follow what

has been instructed by the teacher. In this way of teaching, the teacher serves as the primary source of knowledge and has authority in the classroom (Altun, 2023). Moreover, the teachers use Thai language as the medium of instruction and emphasize more on rote learning and grammar translation. Phoocharoensil (2012) states that students can easily become bored with learning grammar in a traditional way in a second language classroom where the lecture on grammatical rules is the primary focus of the instruction.

To serve the purpose of language teaching, which is to produce learners who can communicate effectively and competently, O'Keeffe et al. (2007) suggest that students need to experience authentic language use based on naturally occurring data rather than ready-made materials available in classrooms. In addition, using corpora in a language classroom aligns with new educational goals outlined in Thailand's Education Reform Law of 1999 which include promoting the learners' literacy, numeracy, language, and technology competency as well as critical thinking and independent, lifelong learning (Kantamara, et al., 2006). By using the computer-assisted analysis of corpus data, the students themselves will have an opportunity to observe and analyze actual grammatical patterns or occurrences employed by native English speakers that are not typically found in most traditional grammar textbooks (Cheng, 2012). In this case, corpus linguistics comes into play in grammar teaching because it describes how grammatical features are used in different contexts (Kheovichai, 2017). According to Choudhury and Rahman (2016), corpus linguistics "has revolutionized the idea of language teaching through deriving rule-patterns from authentic or real-world texts" (p.47). Therefore, the integration of a corpus-based approach into grammar instruction is recommended to foster Thai EFL students' competency for effective communication. To this end, this article aims to provide a fundamental background of grammar instruction and how corpus data can be used in Thai EFL classrooms. Additionally, suggestions are provided regarding the use of corpora for English language teachers.

Importance of Grammar in Language Learning

Grammar refers to the system of sound, structure, and meaning of a language (Afroogh, 2019). According to Huddleston and Pullum (2002), grammar is "the principles or rules governing the form and meaning of words, phrases, clauses, and sentences" (p. 3). Grammar is also recognized as "an essential, inescapable component of language use and language learning" (Han, 2023, p.10). Zhang and Tang (2025) define grammar as "the

cornerstone of English language learning and its importance continues throughout the learning process and its practical application" (p.1).

For second language learners, grammar is an essential factor in language acquisition and plays a predominant role in language learning (Ellis, 2006). Phoocharoensil (2020) insists that a clear understanding of English grammar enhances learners' accuracy in comprehension and production of a second language. In other words, grammar knowledge can assist learners to improve their language proficiency. Phieanchang, (2024) asserts that grammar is important +to language learners in speaking and writing because it contributes to logical thinking and precision in expression. Using grammar structures in a language correctly is necessary for learners in terms of avoiding confusion and ensuring clarity in communication.

According to Oktavianti et al. (2023), grammar is useful and essential since it is strongly related to other language skills. In terms of speaking and listening, the knowledge of grammar helps speakers and listeners better understand communication. Moreover, learners with correct grammar can speak confidently and fit in social, academic, or business settings successfully. This increases opportunities to work in leadership positions where communication is essential (Shivam, 2025). For writing and reading, learners with a clear understanding of grammar rules can write more precise sentences and paragraphs. Although they have limited vocabulary, grammar knowledge aids them to learn English faster since they are able to apply grammar structures to summarize their idea in a writing task (Phuwarat & Boonchukusol, 2020). Furthermore, grammar assists readers to comprehend reading texts more effectively and allows them to appreciate different writing styles and genres more deeply (Shivam, 2025). Afroogh (2019) adds that successful reading comprehension is dependent on grammar knowledge of the readers. To sum up, without good grammar, it is almost impossible to achieve clear communication.

English Grammar Instruction

Apparently, grammar teaching is necessary for learners of the English language. The aim of learning English grammar structures is to understand and obtain grammatical knowledge in terms of its form, meaning, and use in context to express meaning correctly (Zhang & Tang, 2025). For language teachers, the purpose of teaching grammar is to enhance their students' use of grammatical features accurately and fluently in communication (Ellis, 2014). Gündüz

(2022) points out that grammar instruction can help develop a second language in terms of supporting the correct use of language structures and providing effective communication.

Many language teachers have employed grammar-translation approach or teacher-centered education in their grammar teaching. For this approach, the teacher primarily focuses on explaining the grammar rules and providing drills and practice to students through the memorization process, which involves translating and repeating information. Marsden (2005) states that traditional grammar teaching is probably most suitable for academically-oriented students. However, it has not yet been demonstrated to achieve the objectives mentioned in the communicative approach. A major drawback of this method is that the students do not have an opportunity to develop their critical thinking and problem-solving skills. Furthermore, they are not encouraged to tailor their own learning needs, leading to a lack of students' autonomy, engagement, and motivation in language learning (Altun, 2023).

Due to the failure of the conventional grammar teaching method, researchers and practitioners have attempted to redesign grammar instruction to be more innovative and effective. Grammar instruction has been developed in communicative language classrooms over the years. Until recently, grammar instruction could be delivered in a language classroom through different perspectives. Larsen-Freeman (2003) views grammar teaching as a more holistic approach, emphasizing the interconnectedness of three essential dimensions: form (syntax), meaning (semantics), and use (pragmatics). It is believed that the use of three-dimensional grammar theory can effectively promote students' comprehensive language performance (Zhang & Tang, 2025). According to Ellis (2014), grammar can be delivered in two principled ways: explicitly and implicitly. Explicit grammar instruction focuses on direct explanations of structures and rules. Drills, exercises, and traditional textbooks are commonly used in the classroom. Implicit instruction refers to learning through exposure to language and practice, rather than direct instruction. Authentic materials and communicative activities are normally provided for the students. This approach helps promote natural language acquisition and competence.

In addition, grammar can be taught deductively and inductively (Nunan, 2005). Deductive teaching is regarded as a traditional, teacher-centered instruction. It is also known as rule-driven teaching (Sukavatee, 2023). That is to say, grammatical rules are explained explicitly at the beginning of the lesson, followed by examples related to those rules. Then, students practice the rules by doing language activities. Although the deductive method saves

time, and is easy to process and develop students' cognitive skills, it can also lessen students' involvement and interaction. Additionally, students must rely heavily on their memory (Shirav & Nagai, 2022). On the contrary, inductive teaching focuses on providing students with examples of grammatical rules implicitly. Then, the students are encouraged to think critically, based on their previous language knowledge, to analyze the given examples and formulate a rule on their own. According to Widodo (2006), inductive grammar teaching promotes learning autonomy, enhances motivation, and develops cognitive and problem-solving skills. However, it is time-consuming and requires strategic lesson planning. Moreover, students may develop incorrect grammatical patterns.

In conclusion, a language teacher should integrate adaptive learning technologies that tailor grammar exercises to meet the learners' needs, allowing for personalized and efficient practice. Additionally, language learning activities should be designed based on real-world tasks to encourage students to apply grammar rules in authentic contexts and foster practical language use (Han, 2023). Oktavianti et al. (2023) add that it is essential to provide appropriate authentic materials for students when grammar is taught. Moreover, the teacher should employ a variety of teaching methods to encourage the interests and learning process of the students (Altun, 2023). With these combined, a teacher can promote the effectiveness of grammar teaching and enhance a more engaging and tailored learning experience for students.

Corpus Linguistics and Corpora

Corpus linguistics is an approach for language investigation that emphasizes the analysis of real samples of language use. Bennett (2010) states that the main purposes of corpus linguistics are to investigate particular patterns that are associated with lexical or grammatical features and to examine the way in which these patterns differ within varieties and registers. Moreover, Staples and Anthony (2023) assert that corpus linguistics can be regarded as a research methodology assisting researchers, curriculum developers, teachers, and learners to understand language use in different registers, such as journal articles and conversation, by analyzing a large set of authentic texts, called a corpus (or corpora in a plural form).

O'Keeffe et al. (2007) define a corpus as "a collection of texts, written or spoken, which is stored on a computer" (p.1), which is used for qualitative and quantitative analysis

of language features (Choudhury & Rahman, 2016). According to Huang and Yao (2015), a corpus is defined as “a collection of examples of language in use that are selected and compiled in a principled way and corpus linguistics as linguistic studies of such corpora” (p.949). Bennett (2010) gives a definition of a corpus as “a principled collection of authentic texts stored electronically that can be used to discover information about language that may not have been noticed through intuition alone” (p.12). Similarly, Crystal (1992) defines a corpus as “a collection of linguistic data either compiled as written texts or as a transcription of recorded speech” (p.85). A corpus has been considered as one of the most effective, precise, and up-to-date tools in applied linguistic research (Cheng, 2012; Phoocharoensil, 2012).

The emergence of corpus linguistics and corpora has changed everything. This is due to the power of computers to analyze huge numbers of written and spoken texts, allowing us to observe how English grammar has evolved in authentic situations (McCarthy, 2019). In other words, computer technology has made it possible to store and analyze complex patterns of language structures from a large volume of data available online, representing natural language use (Kwon, 2004; Lee, 2011). To study these language features in naturally occurring discourse allows researchers to be able to “identify systematic differences in the functional use of each variant” (Biber, 2005, p.1).

Currently, various online corpora have been created and are easily accessible on the web for teachers and learners of English to use as a resource for learning grammatical structures. Examples of well-known corpora include the Corpus of Contemporary American English (COCA), the Corpus of Historical American English (COHA), the Brown Corpus, the British National Corpus (BNC), the International Corpus of English (ICE), the Michigan Corpus of Academic Spoken English (MICASE), and the TIME Magazine Corpus.

In a language classroom, corpora data can be utilized directly and indirectly (Khemkullanat & Khongput, 2023). The direct use of corpora refers to the situation where learners are provided an opportunity to directly access the corpus data and figure out meaning, lexical, and grammatical rules on their own. In contrast, indirect use of corpora means learning from ready-made materials modified or prepared by the teacher in advance. These materials are called corpus-informed materials, meaning they are register-based and context-appropriate (Oktavianti et al., 2023). To increase the effectiveness of grammar teaching, teachers are recommended to implement both direct and indirect methods in corpus-based instruction. Teaching grammar using corpus-based data enables students to

discover grammatical patterns or occurrences that are not typically found in most traditional grammar textbooks (Cheng, 2012), as well as to enhance their inductive learning and self-discovery skills (Satchayad & Charubusp, 2023).

Corpora used in the language classroom can be categorized into four major types: generalized, specialized, learner, and pedagogic corpora (Bennett, 2010). The teachers can choose corpora based on the purpose of the corpus and the learning objective. Generalized corpora are very large, containing more than 10 million words across a variety of written (e.g., newspapers, magazines, fiction, nonfiction) and spoken genres (conversations, business meetings, government proceedings). Generalized corpora are created to represent all aspects of a language, and the findings can be generalized to other similar contexts. Specialized corpora are constructed to be a representative of language produced by particular user groups, such as children and teenagers, or in a specific setting, such as a university or business. Learner corpora are collections of written or spoken language data produced by language learners. The aim of this corpora is to study second language acquisition (SLA) and provide information on language used by students in terms of language patterns, errors, and the influence of native language. Pedagogic corpora can be applied as a resource for language teaching and research and as a self-reflective tool for teacher development. Moreover, they can be used to examine teacher-student dynamics through interaction in the classroom.

Corpus-based Approach in Grammar Instruction in Thai EFL Contexts

More recently, EFL teachers have begun to apply authentic data available from corpus linguistics in their classrooms to enhance learners' grammar competency. To do this, students will have a chance to be exposed to grammatical patterns from large text collections representing language use in real-life situations (Biber, 2005). The implementation of the corpus approach in grammar teaching in language classrooms is considered an inductive teaching method that promotes learner-centeredness (Phoocharoensil, 2012; Samburskiy, 2014) by employing the fine-grained information from corpus data (Klomkaew & Boontam, 2023).

Previous studies conducted in Thai EFL contexts have revealed the success of corpus linguistics implemented for improving learners' grammatical knowledge in various aspects and at different educational levels. Suriyapee and Pongpairoj (2022) examined the effects of a corpus-based teaching approach using concordance lines in the Corpus of Contemporary

American English (COCA) on the second language acquisition of English *to-infinitive* and *gerund* complements among low-English-proficiency young Thai learners of English and found out their attitudes toward corpus-based instruction. The results showed that the students' knowledge of the *to-infinitive* and *gerund* complement increased after five weeks of corpus instruction. Moreover, the students reported that they could see authentic examples by referring to the corpus instead of memorizing the verbal complements. By using information available from COCA, they were able to solve the language problem themselves with confidence, pride, and motivation.

Jantarabang and Tachom (2021) investigated the effectiveness of a corpus-based approach on students' grammatical competence for writing narrative paragraphs and to elicit Thai EFL high school students' perceptions towards using corpus-based instructional materials to improve the students' paragraph writing skills. The findings revealed that the overall number of grammatical errors in the post-writing test was lower than that of the pre-test. This suggests that the corpus-based approach can help enhance EFL students' development of their writing ability. Furthermore, the students had positive attitudes towards the use of corpus material to enhance their grammar competence because they had more exposure to the use of corpus data through writing tasks for several weeks.

Dewiloh (2022) implemented the Data-Driven Learning (DDL) method to investigate the effectiveness of Thai EFL secondary students' knowledge of English phrasal verbs. Moreover, the relationship between students' grade level and their learning performance through DDL, as well as their attitudes towards learning through DDL, was also examined. The results indicated that DDL improved students' knowledge of English phrasal verbs; however, there was no correlation between students' educational level and their learning ability through DDL. Furthermore, the students' attitudes towards learning grammar through DDL were not positive because they reported difficulty in reading and understanding the concordance lines. The possible reason for this phenomenon is that the students might prefer a teacher-centered method rather than a learner-centered method, such as DDL.

Khemkullanat and Khongput (2023) employed a corpus-assisted approach with Data Driven Learning (DDL) to foster collocational knowledge of Thai EFL university students. The results revealed that students' knowledge of collocations in all patterns improved (adjective-preposition collocation, noun-preposition collocation, and verb-preposition collocation, respectively) after learning through a corpus-based approach with DDL. Additionally, the

students were aware of the grammatical accuracy of collocational patterns, particularly the correct forms of collocations. They realized that the correct use of a collocation is significant in communication and valued collocations as a mechanism helping them to become native-like English writers.

At a sentence level, Klomkaew and Boontam (2023) investigated the impacts of paper-based Data-Driven Learning (DDL) activities based on the data available from the Corpus of Contemporary American English (COCA) on the grammar learning of conditional sentences of Thai EFL undergraduate students majoring in English. Plus, their attitudes towards learning through the DDL approach were examined. The findings demonstrated that learners' grammatical knowledge related to conditional sentences improved after the DDL teaching method. Apart from that, the students showed positive attitudes towards DDL learning activities. They considered that DDL is a useful tool supporting them in acquiring grammatical knowledge and language patterns by themselves. Key findings from Thai studies are summarized in Table 1.

Overall, existing studies on the application of the corpus-based approach in Thai EFL classrooms at various educational levels have yielded positive results in terms of assisting learners in overcoming problems when encountering difficult grammatical features, such as English *to-infinitive* and *gerund* complements, phrasal verbs, collocations, and conditional sentences. Additionally, the students found the corpus-based approach helpful and different from the traditional grammar teaching method. That is, it promotes autonomous learning since it offers a good experience to discover patterns of grammatical structures themselves. However, a study by Dewiloh (2022) showed a contrasting result. Thai secondary students in the study preferred learning grammar in a conventional way rather than a corpus-based approach. The possible reason is that they did not want to explore the knowledge on their own because they were worried about making a mistake or drawing the wrong conclusions. Therefore, they heavily relied on the teacher as the major source of knowledge in the classroom. Based on this finding, it is worth noting that the corpus-based teaching approach, which is considered an inductive teaching method, may not be suitable for all Thai students. Regarding the research methodology employed in previous studies, learners' perceptions of using corpora in grammar instruction were primarily collected through questionnaires and surveys. Further studies should employ additional qualitative methods, such as interviews and observation, to verify the results and increase the credibility and validity of the research.

Table 1 Key Findings from Thai Studies on the Use of Corpora

Authors	Participants	Corpus Used	Outcomes
Suriyapee & Pongpairoj (2022)	Young Thai learners of English with low English proficiency	Concordance lines on <i>to-infinitive</i> and <i>gerund</i> complements	<ul style="list-style-type: none"> - The students' knowledge of the <i>to-infinitive</i> and <i>gerund</i> complement increased after five weeks of the corpus instruction. - The students were able to solve the language problem themselves with confidence, pride, and motivation.
Jantarabang & Tachom (2021)	Thai EFL high school students	A corpus-based instructional material used to improve the students' paragraph writing skills	<ul style="list-style-type: none"> - Grammatical errors in the post-writing test was lower than that of the pre-test. - The students had positive attitudes towards the use of corpus material to enhance their grammar competence.
Dewiloh (2022)	Thai EFL secondary students	Data-Driven Learning (DDL) method employed to increase phrasal verbs knowledge	<ul style="list-style-type: none"> - DDL improved the students' knowledge of English phrasal verbs. - There was no correlation between students' educational level and their learning ability through DDL. - The students' attitudes regarding learning through the DDL were not positive.
Khemkullanat & Khongput (2023)	Thai EFL university students	Data Driven Learning (DDL) used to foster collocational knowledge	<ul style="list-style-type: none"> - The students' knowledge of collocations in all patterns improved (adjective-preposition collocation, noun-preposition collocation, and verb-preposition collocation, respectively) after learning through a corpus-based approach with DDL. - The students were aware of the grammatical accuracy of collocational patterns.

Authors	Participants	Corpus Used	Outcomes
Klomkaew & Boontam (2023)	Thai EFL undergraduate students majoring in English	Data-Driven Learning (DDL) activities implemented to improve conditional sentences knowledge	- Learners' grammatical knowledge related to conditional sentences improved after the DDL teaching method. - The students showed positive attitudes towards DDL learning activities.

Benefits of Corpora in English Grammar Instruction

According to Choudhury and Rahman (2016), corpus linguistics has positively influenced the teaching of grammar. One of the most important advantages of corpora is their authenticity. Corpora provide authentic and real-life examples which expose students to grammatical patterns used in natural settings (Reppen, 2011; Suriyapee & Pongpairoj, 2022). Those are different from the language that appears in a textbook that often displays unrealistic language for real-life communication (Satchayad & Charubusp, 2023). For example, based on corpus data drawn from COCA, the construction of *get-passive* was more generally used in natural spoken English when compared to the *be-passive* structure, which is normally found in the textbooks (Phoocharoensil, 2020). Therefore, it is suggested that teachers should not rely solely on textbooks and dictionaries. Instead, they are supposed to develop instructional materials based on authentic and naturally occurring language use.

Another advantage of corpora deals with the representativeness. With a huge amount of language data, corpora can represent linguistic features that appear in the corpus. According to Biber (1993), representativeness refers to "the extent to which a sample includes the full range of variability in a population" (p.243). For example, the British National Corpus (BNC), a large collection of written and spoken language samples, is designed to represent a wide range of British English from the late twentieth century. Therefore, the BNC is considered a valuable resource for studying contemporary British English, as it can be generalized to a larger linguistic population. McEnery and Wilson (2001) point out that representativeness is an essential property in corpus-based material design. As Biber (2005) states, grammar-related material developers and designers who use resources and techniques available from corpus linguistics put greater emphasis on the representativeness of the database.

Corpus linguistics can also promote learner autonomy and independent learning. Students are allowed to observe grammar rules and structures on their own by investigating actual examples of language use in different situations through corpus-based materials and activities in the classroom (Çiftci & Özcan, 2021). Likewise, they can get access to online information on how people use a particular grammar point in real-life situations outside the classroom by themselves. For instance, students who are unclear about the use of the words “effect” and “affect” might consult a corpus tool to examine the actual sentences that contain each of them (Sinclair, 2004). This way of learning enables students to take charge of their own education and promotes critical thinking ability and autonomous linguistic research.

Challenges of Using Corpora in English Grammar Instruction

Corpora have been widely acknowledged as a valuable resource and an innovative teaching tool in language classrooms. However, there are some challenges that teachers and learners need to be aware of when using corpora. A major challenge is the selection of the corpus. A particular corpus is always designed for a specific purpose, depending on its intended purpose (Boumediène, 2016). The teacher should carefully select a corpus that is appropriate for a specific teaching context, meets the learners’ needs and analytical abilities, represents a target register, and aligns with the class’s learning objectives (Khemkullanat & Khongput, 2023; Suriyapee & Pongpairoj, 2022). For example, specialized corpora compiled from professional or academic publications within a particular discipline are suitable for examining the use of vocabulary and grammar in specific fields, such as medicine, law, and engineering. These corpora should be chosen as resources for the English for Specific Purposes (ESP) class because they can represent authentic use of domain-specific language structures, specialized jargon, and terminologies (Bowker & Pearson, 2002). In this case, the teacher can ensure that the corpus truly reflects a broad range of language users (Dalieva, 2024).

Another challenge is that corpus resource tools need specific training sessions to use them effectively (Alamri & Arabia, 2022). Teachers may need guidance in order to implement a corpus-based approach effectively in their actual teaching practices (Çalışkan & Gönen, 2018). Samburskiy (2014) points out that language teachers are often unwilling to integrate corpus-based tasks into their classes because they lack understanding of how to use them properly. Also, they may lack awareness of how to design corpus-based teaching materials and activities. The students should also be trained to familiarize themselves with corpus consultation under

the guidance of the teachers. According to the research conducted by Elsherbini and Ali (2017), many students experienced difficulties when logging in to the corpus website, such as COCA, since it requires proper search techniques to get data and interpret corpus findings. Due to their unfamiliarity with corpus investigation, learners have to spend a lot of time trying to understand the complexity of the grammar rules. Thus, the teachers should be trained to use the corpus efficiently and scaffold students to use various features of the corpus. During the training, participants should be provided with modeling, tutorials, lessons, handouts or booklets, and multimedia digital files to ensure a comprehensive understanding of corpus investigation.

Language proficiency of the students is also a challenge in implementing a corpus-based teaching approach in a grammar lesson. Students with high proficiency levels are normally more advantageous in terms of designing their own corpus investigation since they have the freedom to learn particular language aspects which they are interested in (Bernadini, 2000). Elsherbini and Ali (2017) note that advanced learners benefit more from large corpora, whereas lower-level students often face a lot of difficulties in direct use of corpora because they feel overwhelmed with too many corpus data examples and do not know how to deal with them properly. For this problem, low proficient learners should begin with smaller corpora, which are limited in size but are easier to access. After that, they can move to larger corpora. Guidance from the teacher can help students manage a large amount of corpus data so that they can gradually develop a better understanding of corpus use (Charles, 2007).

Suggestions

Corpus linguistics plays a part in reforming grammar instruction in Thai EFL contexts. To better enhance efficiency in teaching grammar in order to foster learners' grammar knowledge, specific steps for implementing a corpus-based approach in classrooms are suggested for the teachers. Before the instruction, the teachers should begin with identifying learners' specific needs of grammatical points through a needs analysis which helps understand their areas of problems in grammar learning. After that, the teachers should carefully select a corpus from a reliable source that is fit to learners' needs and learning objectives. Then, they prepare teaching materials based on the selected corpus data. In this step, teachers should ensure that the selected data can accurately represent the linguistic structures that appear in the corpus. During instruction, teachers are suggested to act as

directors and coordinators of learner-initiated linguistic research (John, 1986). In other words, teachers should approach their students as language researchers by allowing students to access linguistic data independently. In this way, grammatical rules are delivered through an inductive approach, i.e., the learners analyze and discover the rule on their own. Additionally, Khemkullnat and Khongput (2023) recommend that teachers provide an explanation of corpus-based exercises to their students whether the grammatical patterns they have noticed are correct. When the students face problems when using a corpus for grammar learning, they should be provided with strategies or solutions to overcome them. After the instruction, the teachers should have the students do corpus-based exercises and activities to reinforce learning and facilitate the retention of new grammar knowledge. Moreover, the students should be asked to take a test in order to assess the correct use of grammatical patterns learned in the classroom, as well as to examine their progress after the grammar lesson. The framework for implementing a corpus-based approach in grammar instruction is illustrated in Figure 1.

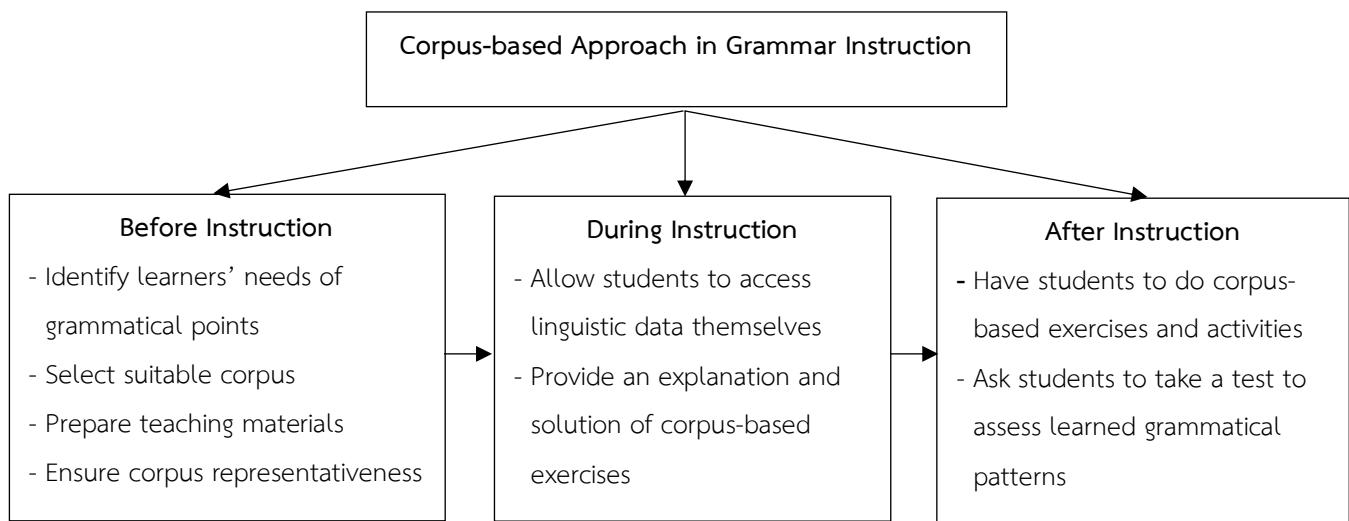


Figure 1. Steps for Implementing Corpus-based Approach in Grammar Instruction

Conclusion

Grammar plays a crucial role in second language learning and teaching. It is believed to be the most difficult and complex component for Thai EFL students because it consists of a list of rules that learners should achieve to be fluent in the target language (Oktavianti et al., 2023). Also, it is the foundation of the language. Without knowledge of grammar, students

are unable to understand a second language (Phuwarat & Boonchukusol, 2020). Grammar has traditionally been taught conventionally in Thai EFL classrooms. Until recently, corpus linguistics has shed new light on several aspects of English grammar via the analysis of authentic examples in real-life settings (Kheovichai, 2017). Through a corpus-based approach, teachers and learners of English can move beyond rote memorization and transform grammar teaching into a dynamic study by harnessing the power of digital language analysis (Eragamreddy, 2024). Biber (2006) concludes that corpus-based analysis, when implemented in language pedagogy, significantly contributes to the study of authentic patterns of language use and grammatical features found in spoken and written language in real-life, everyday communication.

References

Afroogh, M. R. (2019). Grammar is the heart and basis of language teaching and learning: English grammar and its role in ELT. *Annals of Language and Literature*, 3(1), 6-11. <https://doi.org/10.22259/2637-5869.0301002>

Alamri, B., & Arabia, S. (2022). The role of corpus linguistics in grammar instruction: A review of literature. *International Journal of Linguistics*, 14(6), 158-166. <https://doi.org/10.5296/jjl.v14i6.20500>

Altun, M. (2023). The ongoing debate over teacher centered education and student centered education. *International Journal of Social Sciences & Educational Studies*, 10(1), 106-110. <https://www.ijsses.tiu.edu.iq/index.php/ijsses/article/view/128/119>

Azar, B. (2007). Grammar-based teaching: A practitioner's perspective. *TESL-EJ*, 11(2), 1-12. <https://www.tesl-ej.org/ej42/a1.pdf>

Bennett, G. (2010). *An introduction to corpus linguistics: Using corpora in language learning classroom: Corpus linguistics for teachers*. University of Michigan Press.

Bernadini, S. (2000). *Competence, capacity, corpora*. CLUEB.

Biber, D. (1993). Representativeness in corpus design. *Literary and Linguistic Computing*, 8(4), 243–257. <https://doi.org/10.1093/lrc/8.4.243>

Biber, D. (2005). Corpus linguistics and the study of English grammar. *Indonesian Journal of English Language Teaching*, 1(1), 1-22. <https://ejournal.atmajaya.ac.id/index.php/ijelt/article/view/1398/683>

Biber, D. (2006). Parsing and grammar description, corpus-based. In K. Brown (Ed.), *Encyclopedia of Language & Linguistics* (2nd ed., pp. 197-205). Elsevier.

Boumediène, B. (2016). Corpus linguistics: Applications, implications, and limitations. *Revue Maghrébine des Langues*, 10(1), 185-205. <https://asjp.cerist.dz/en/article/135079>

Bowker, L., & Pearson, J. (2002). *Working with specialized language: A practical guide to using corpora*. Routledge.

Çalışkan, G., & Gönen, S. İ. K. (2018). Training teachers on corpus-based language pedagogy: Perceptions on using concordance lines in vocabulary instruction. *Journal of Language and Linguistic Studies*, 14(4), 190-210. <https://www.jlls.org/index.php/jlls/article/view/1020/430>

Charles, M. (2007). Reconciling top-down and bottom-up approaches to graduate writing: Using a corpus to teach rhetorical functions. *Journal of English for Academic Purposes*, 6(4), 289-302. <https://doi.org/10.1016/j.jeap.2007.09.009>

Cheng, W. (2012). *Exploring corpus linguistics: Language in action*. Routledge.

Choudhury, A. S., & Rahman, M. (2016). Using corpora in grammar pedagogy: An exploration of some central issues. *Intellection: A Bi-annual Interdisciplinary Research Journal*, 4(2), 46-53.

Çiftci, H., & Özcan, M. (2021). A contrastive analysis of traditional grammar translation method and communicative language teaching in teaching English grammar and vocabulary. *International Online Journal of Education and Teaching (IOJET)*, 8(2), 709-729. <https://files.eric.ed.gov/fulltext/EJ1294310.pdf>

Crystal, D. (1992). *An Encyclopedic Dictionary of Language and Languages*. Blackwell.

Dalieva, M. (2024). Methods, challenges, and ethical considerations in data collection of corpus compilation. *Innovative Technologica: Methodical Research Journal*, 3(3), 1-6. <https://doi.org/10.47134/innovative.v3i3.122>

Dewiloh, K. (2022). The effects of data-driven learning (DDL) on the teaching of English phrasal verbs to Thai EFL secondary students (Publication No. 25656421042117LKS) [Master's thesis, Thammasat University]. https://digital.library.tu.ac.th/tu_dc/frontend/info/item/dc:307017

Ellis, R. (2006). Current issues in the teaching of grammar: An SLA perspective. *TESOL Quarterly*, 40(1), 83-107. <https://doi.org/10.2307/40264512>

Ellis, R. (2014). Grammar teaching for grammar learning. *Babylonia*, (2), 10-15. http://babylonia.ch/fileadmin/user_upload/documents/2014-2/Ellis.pdf

Elsherbini, S. A., & Ali, A. D. (2017). The effects of corpus-based activities on EFL university students' grammar and vocabulary and their attitudes toward corpus. *Journal of Research in Curriculum Instruction and Educational Technology*, 3(1), 133-161.
<https://doi.org/10.21608/jrciet.2017.24482>

Eragamreddy, N. (2024). Grammar instruction in the digital age: The impact of corpus linguistics. *International Journal of English Literature and Social Sciences*, 9(2), 161-170. <https://doi.org/10.22161/ijels.92.24>

Gündüz, F. (2022, July 7). *The value of grammar instruction in L2 development*. Medium. <https://medium.com/@fatmaggunduz/the-value-of-grammar-instruction-in-l2-development-ca2bb795623b>

Han, J. (2023). On approaches to English grammar teaching and its current situation. *Open Access Library Journal*, 10, 1-12. <https://doi.org/10.4236/oalib.1111059>

Huang, C., & Yao, Y. (2015). Corpus linguistics. In J. D. Wright (Ed.), *International Encyclopedia of the Social & Behavioral Sciences* (2nd ed., pp. 949-953). Elsevier.

Huddleston, R. & Pullum, G.K. (2002). *The Cambridge grammar of the English language*. Cambridge University Press.

Hymes, D. H. (1972). On communicative competence. In J. B. Pride & J. Holmes (Eds.), *Sociolinguistics: Selected Readings* (pp. 269-293). Penguin.

Jantarabang, N., & Tachom, K. (2021). An investigation of students' experiences with corpus-based approach on high school students' English paragraph writing. *Journal of SaengKhomKham Buddhist Studies*, 6(1), 1-15. <https://so02.tci-thaijo.org/index.php/jsbs/article/view/242694>

John, T. (1986). Micro-concord: A language learner's research tool. *System*, 14(2), 151-162. [https://doi.org/10.1016/0346-251X\(86\)90004-7](https://doi.org/10.1016/0346-251X(86)90004-7)

Kampookaew, P. (2020). An analysis of grammatical errors made by Thai EFL university students in an EAP writing class: Issues and recommendations. *rEFLections*, 27(2), 246-273. <https://doi.org/10.61508/refl.v27i2.248862>

Kantamara, P., Hallinger, P., & Jatiket, M. (2006). Scaling-up educational reform in Thailand: Context, collaboration, networks, and change. *Planning and Changing*, 37(1), 5-23. <https://files.eric.ed.gov/fulltext/EJ756213.pdf>

Kaosayapandhu, M. (2023). Application for grammar skills: A case study of Thai EFL undergraduates. *PASAA Journal*, 67 (1), 309-329. <https://doi.org/10.58837/CHULA.PASAA.67.1.10>

Khemkullanat, P., & Khongput, S. (2023). The effectiveness of corpus-assisted approach in learning grammatical collocations of Thai undergraduate students in an EFL classroom. *rEFLections*, 30(3), 913–939. <https://doi.org/10.61508/refl.v30i3.268920>

Kheovichai, B. (2017). A Corpus-based analysis of BE + Being + Adjective in English from the appraisal framework perspective. *Humanities, Arts and Social Sciences Studies*, 17(2), 97-132. <https://doi.org/10.14456/sujsha.2017.1>

Klomkaew, T., & Boontam, P. (2023). Effectiveness of grammar learning through data-driven learning (DDL) with undergraduate students. *Shanlax International Journal of Education*, 11(4), 30-44. <https://doi.org/10.34293/education.v11i4.6385>

Kwon, H. (2004). Using corpora for studying English grammar. *Korean Journal of English Language and Linguistics*, 4(1), 61-81. G704-001472.2004.4.1.003

Larsen-Freeman, D. (2003). *Teaching language: From grammar to grammaring*. Thomson/Heinle.

Lee, S. (2011). Challenges of using corpora in language teaching and learning: Implications for secondary education. *Linguistic Research* 28(1), 159-178. <https://doi.org/10.17250/khisli.28.1.201104.009>

Marsden, M. (2005) Input-based grammar pedagogy: A comparison of two possibilities. *Language Learning Journal*, 31(1), 9-20. <https://doi.org/10.1080/09571730585200041>

McEnery, T., & Wilson, A. (2001). *Corpus linguistics*. Edinburgh University Press.

McCarthy, M. (2019). *Corpora and grammar*. Cambridge University Press.

Nunan, D. (2005). *Practical English language teaching*. McGraw-Hill.

O'Keeffe, A., McCarthy, M., & Carter, R. (2007). *From Corpus to Classroom: Language Use and Language Teaching*. Cambridge University Press.

Oktavianti, I. N., Eriani, E., Rolyna, I., & Prayogi, I. (2023). Investigating the use of corpus-informed grammar materials in Indonesian EFL classrooms. *Indonesian Journal of English Language Teaching and Applied Linguistics*, 7(2), 417-438. <http://dx.doi.org/10.21093/ijeltal.v7i2.1415>

Phieanchang, T. (2024). Using ChatGPT to improve English grammar skills for EFL learners. *Dhonburi Rajabhat University Journal*, 18(1), 223–240. <https://so02.tci-thaijo.org/index.php/journaldru/article/view/268054>

Phoocharoensil, S. (2012). Language corpora for EFL teachers: An exploration of English grammar through concordance lines. *Social and Behavioral Sciences*, 64, 507-514. <https://doi.org/10.1016/j.sbspro.2012.11.060>

Phoocharoensil, S. (2020). English *get*-passives: Reassessing the frequencies across genres. *GEMA Online Journal of Language Studies*, 20(3), 123-135.
<https://doi.org/10.17576/gema-2020-2003-08>

Phoocharoensil, S. (2023). Teaching English conditionals through Data-Driven Learning (DDL): Perspectives of in-service EFL teachers. *Theory and Practice in Language Studies*, 13(10), 2620-2628. <https://doi.org/10.17507/tpls.1310.21>

Phuwarat, C., & Boonchukusol, N. (2020). The role of grammar teaching in ESL writing. *Sripatum Chonburi Academic Journal*, 16(4), 1-9. <https://so05.tci-thaijo.org/index.php/SPUCJ/article/view/244110>

Reppen, R. (2011). Using corpora in the language classroom. *ELT Journal*, 65(3), 348–350. <https://doi.org/10.1093/elt/ccr036>

Saengboon, S., Kosin, P., & Toomaneejinda, A. (2022). The roles of grammar in English language teaching: Local viewpoints. *PASAA Journal*, 63(1), 179-204. <https://doi.org/10.58837/CHULA.PASAA.63.1.7>

Samburskiy, D. (2014). Corpus-informed pedagogical grammar of English: Pros and cons. *Social and Behavioral Sciences*, 154, 263-267. <https://doi.org/10.1016/j.sbspro.2014.10.148>

Satchayad, P., & Charubusp, S. (2023). The impacts of blended corpus-based instruction on enhancing writing proficiency of Thai university students. *English Language Teaching Educational Journal*, 6(2), 109–123. <https://doi.org/10.12928/eltej.v6i2.7872>

Sermsook, K., Liamnimitr, J., & Pochakorn, R. (2017). An analysis of errors in written English sentences: A case study of Thai EFL students. *English Language Teaching*, 10(3), 101-110. <https://doi.org/10.5539/elt.v10n3p101>

Shirav, A., & Nagai, M. (2022). The effects of deductive and inductive grammar instructions in communicative teaching. *English Language Teaching*, 15(6), 102-123. <https://doi.org/10.5539/elt.v15n6p102>

Shivam, K. (2025, February 1). Importance of grammar in English language. *21K School*. <https://www.21kschool.com/th/blog/the-importance-of-grammar-in-english-language/>

Sinclair, J. (2004). *Trust the text: Language, corpus, and discourse*. Routledge.

Staples, S., & Anthony, L. (2023). Corpus linguistics in EFL language teaching: Insights from research and practice. *JALT Journal*, 45(2), 246-263. <https://doi.org/10.37546/JALTJJ45.2-4>

Sukavatee, P. (2023). The effects of inductive and deductive instruction in a flipped classroom on English grammar ability of undergraduate students in Thailand. *PASAA Journal*, 67(1), 271-308. <https://doi.org/10.58837/CHULA.PASAA.67.1.9>

Suriyapee, P., & Pongpairoj, N. (2022). Corpus to enhance verbal complements among low English proficiency Thai learners. *PASAA Journal*, 63(1), 205-231. <https://doi.org/10.58837/CHULA.PASAA.63.1.8>

Tantiwich, K., & Sinwongsuwan, K. (2021). Thai university students' problems of language use in English conversation. *LEARN Journal: Language Education and Acquisition Research Network*, 14(2), 598-626. <https://so04.tci-thaijo.org/index.php/LEARN/article/view/253282>

Widodo, H. P. (2006). Approaches and procedures for teaching grammar. *English Teaching: Practice and Critique*, 5(1), 122-141. <https://edlinked.soe.waikato.ac.nz/journal/files/etpc/files/2006v5n1nar1.pdf>

Yuh, A. H., & Kaewurai, W. (2021). An investigation of Thai students' English-speaking problems and needs and the implementation collaborative and communicative approaches to enhance students' English-speaking skills. *The Golden Teak: Humanity and Social Science Journal (GTHJ.)*, 27(2), 91-107. <https://so05.tci-thaijo.org/index.php/tgt/article/view/252425>

Zhang, X., & Tang, J. (2025). Literature review on English grammar teaching in light of three-dimensional grammar framework. *Open Access Library Journal*, 12, 1-12. <https://doi.org/10.4236/oalib.1112584>

Citation

Raungsawat, N. (2025). Corpus linguistics and English grammar instruction in Thai educational context. *e-Journal of Education Studies, Burapha University*, 7(2), 1-21.  <https://so01.tci-thaijo.org/index.php/ejes/article/view/280292>