

# Advancing Academic Writing Proficiency: Corpus-Based Implications and Strategies

İsmail Dulkadir,<sup>1</sup> Gökc  Velioglu, Ay  Kaya

*Hacettepe University, Ankara, Türkiye*

---

## Abstract

This research study explores the complex interplay among grammatical structures, style, and vocabulary choices in the context of academic writing. With the goal of improving students' academic writing proficiency, the study presents novel corpus-based tasks designed specifically for English language learners. Three crucial aspects primarily addressed are as follows: the organization of grammatical structures, the use of formal language, and the selection of appropriate words. In addition, the study equips educators with useful corpus-based teaching methodologies that match changing demands in academic writing education. Students are immersed in an interactive learning environment through real-world examples and practical exercises, facilitating their refinement of the passive voice production skills and develop a deep understanding of the intricacies of the English language. Using several corpora such as the British National Corpus (BNC); the Corpus of Contemporary American English (COCA) and software like AntConc for in-depth analysis, the project also introduces a corpus-based application for teaching learners how to properly utilize articles in English. Furthermore, it offers a useful vocabulary-building exercise using learner corpora that prioritize contextual appropriateness and lexical sophistication. Additionally, the study supports the use of corpus analysis to encourage formality in academic writing and suggests group projects to help students become more conscious of formal phrases. This all-encompassing method surpasses traditional academic writing requirements by giving students a profound command of language. The study concludes by highlighting the revolutionary potential of corpus-based strategies in improving academic writing skills and providing insightful information to both teachers and language learners.

**Keywords:** *Corpus, Academic Writing, Formality*

*Received Date: November 1<sup>st</sup>, 2023. Acceptance Date: December 12<sup>th</sup>, 2023.*

  Journal of Language and Education Review. All rights reserved.

---

## 1. Introduction

Gaining proficiency in academic writing involves navigating an intricate path through the subtleties of grammatical intricacies, style, and vocabulary choice. This paper explores the complex issues surrounding academic writing and presents innovative solutions through carefully planned corpus-based exercises created especially for English language learners. The current study's main goal is to equip students with the tools they need to improve their academic writing by focusing on three key areas: grammatical structural organization, formal language use, and appropriate word choice. Concurrently, this project seeks to provide teachers with useful and effective corpus-based teaching strategies, bringing pedagogy in line with the changing needs of academic writing. The subsequent sections include various facets of academic writing, and they reveal creative approaches and useful knowledge that have the potential to transform language education. As each aspect is examined, the focus is not just on resolving issues but also on developing a deep, complex linguistic ability that goes beyond accepted standards. The goal of this all-encompassing strategy is to develop a generation of writers who can meet and even surpass the demands of academic writing.

The purpose of this paper is to address several aspects of academic writing which include the following three dimensions: the style, the use of vocabulary and the structure. Corpus-based tasks that can be implemented by English language teachers in academic writing classes are presented as a solution to relevant problems respectively. We aim to aid the students in improving their academic writing skills in the above-mentioned three dimensions as well as providing the teachers with corpus-based task ideas that can be used in the classroom environment for this purpose.

---

\*ADDRESS FOR CORRESPONDENCE: İsmail Dulkadir Department of Foreign Languages, Hacettepe University, E-mail address: [ismail.dlkdroglu@gmail.com](mailto:ismail.dlkdroglu@gmail.com) ORCID ID: 0009-0002-9878-6082

G c  Velioglu Department of Foreign Languages, Hacettepe University ORCID ID:0009-0005-5437-0124  
Ay  Kaya Department of Foreign Languages, Hacettepe University ORCID ID: 0009-0006-7581-6428

## 1.1. Corpus / Corpora

Within the field of linguistics, a corpus is an orderly digital collection that has several genres, languages, and historical periods. This extensive collection provides significant insights on language use in spoken and written communication. It includes literary works, intellectual discourse, and colloquialisms. Corpora, which serve as language databases, are vital resources that let researchers look into obscure topics that are outside the purview of traditional methods. Their diversity enables academics to identify subtle patterns, trends, and frequencies in language structures. A corpus breaks free from conventional limitations and becomes a living, breathing ecosystem where language comes together for close inspection.

The study of corpus linguistics and language education has grown in importance, offering important new perspectives on pedagogy, curriculum development, and language learning. As a methodological approach, corpus linguistics methodically examines large-scale, electronically recorded text collections to find trends, usage, and language structures. This mutually beneficial relationship is examined, clarifying how corpus-based findings improve language instruction. The dynamic interaction between linguistics and education forms the basis for comprehending how language study is changing and for developing successful language education strategies. The ramifications of this symbiosis will be explored in detail in the following sections, offering a thorough knowledge of its influence on promoting academic writing.

### 1.1. *Corpus and Language Teaching*

As a methodological technique, corpus linguistics has had a major impact on language education by providing valuable insights into pedagogy, curriculum creation, and language acquisition processes. Corpora provide teachers with essential support in their efforts to establish a vibrant and genuine language learning environment. Corpora offer access to a wide range of linguistic features found in real-world language samples from different registers, genres, and situations, going beyond the constraints of traditional textbooks. By incorporating real language into teaching materials, learners can interact with the nuances of language use and develop a more thorough comprehension of language within its context.

With their data-driven approach, corpora play a pivotal role in curriculum design and syllabus preparation in language instruction. Teachers can find high-frequency terms, common collocations, and typical language patterns unique to particular contexts by analyzing corpora. The selection of relevant material is guided by this empirical understanding of language usage, which guarantees that language instruction can be considered related with learners' communicative needs.

Lastly, outside of the classroom, corpora have a positive effect on language instruction by encouraging self-directed learning. Online language resources and corpora enable students to independently investigate and evaluate authentic language material. Through self-directed interaction, language learners are better equipped to negotiate language use in a variety of everyday contexts, promoting linguistic autonomy and overall proficiency.

## 2. Purpose of the Current Study

The purpose of this paper is to address the problems related to academic writing which include the following three dimensions: the style, the use of vocabulary and the grammatical structure. Corpus-based tasks that can be implemented by English language teachers in the academic writing classes are presented as a solution to the related problems respectively. Our aim is to aid the students in improving their academic writing skills in the above-mentioned three dimensions as well as providing the teachers with the corpus-based task ideas that can be used in the classroom environment for this purpose.

## 2. Literature Review

The complex interactions of grammar, vocabulary, and formality greatly influence academic writing. Grammar is a set of rules and conventions that specify how ideas are arranged and conveyed, acting as the structural cornerstone of language. According to Hinkel (2004), grammatical integrity is critical to guaranteeing that a writer's opinions are understandable, independent of intelligence or organizational abilities. According to Baleghizadeh & Gordani (2012), adhering to grammatical rules in academic settings ensures accuracy, clarity, and effective transmission of complicated ideas.

Both vocabulary and grammar are crucial for effective communication. According to To'laganova & Muratovna Panabayeva (2022), the expressive and receptive components of vocabulary serve as the cornerstone of coherent communication. According to Cain and Oakhill (2014), the development of thematic coherence and coherent communication are supported when there is a rich, precise, and detailed semantic knowledge of words. Because the English vocabulary is so large—more than 170,000 words—learning this important ability requires a sophisticated comprehension and developmental method, as noted by Brooks et al. (2021).

Along with vocabulary and grammar, formality becomes a crucial consideration in academic writing, especially for students learning English for Academic Purposes. Formal language is defined by Heylighen and Dewaele (1999, p. 34) as an effort to minimize context-dependence and fuzziness in expressions in order to avoid ambiguity. The importance of formality is highlighted by Larsson et al. (2023), who claim that writing with the necessary formality is crucial for college students, particularly those who hope to publish their work. Because of this emphasis on formality—which is indicated by the lack of overtly casual elements—students must develop language skills that will allow them to successfully negotiate the challenges of formal speech (Liardet et al., 2019). The trio of grammar, vocabulary, and formality emerges as the cornerstone for the development of strong academic writing skills in this holistic viewpoint.

#### **4. Increasing Academic Writing Proficiency: Corpus-Based Strategies for Promoting Grammar Learning**

Proficiency in academic writing necessitates a thorough method that deals with the complexities of grammar acquisition. As the cornerstone of coherent communication, corpus-based solutions stand out as a critical answer to problems related to grammatical nuances. Students can study real-world examples through dynamic and interactive learning made possible by the usage of corpora. Students' understanding of academic writing norms, traditions, and contextual nuances is improved by this immersion. Through the implementation of corpus-based tactics, teachers cultivate a deep comprehension of grammar, hence augmenting students' writing skills. The following parts will look at corpus-based methods for teaching passive voice instruction and English articles, which are important for developing linguistic coherence and accuracy. The following sections will specify the problems faced and cover techniques that use corpus analysis to enhance students' comprehension and provide teachers with useful information in these areas.

#### **5. Refining Academic Writing: Instructing Passive Voice via Corpus Analysis**

##### *5.1 Learning Problem*

As stated before, language grammar has been a crucial component in improving written and spoken communication correctness and in assisting speakers in fixing their errors. It is particularly crucial for individuals who speak English as a second or foreign language to understand English grammar. One important aspect of English grammar that is widely used in formal and academic writing is passive construction (Halik & Kareema, 2020). Learners encounter difficulties when it comes to misforming passive verbs, keeping active order in passive form, missing or misplacing the agent, and confusing active form with passive order.

Hinkel and Fotos (2001) observe that a major obstacle in teaching L2 grammar is conveying the meaning, applications, and functions of passive voice. Although students can correctly transform active sentences into passive ones, there is still a barrier when it comes to successfully integrating passive voice into written communication, indicating the need for focused teaching strategies and interventions.

##### *5.2 Suggested Task*

This extensive task, which is tailored for second language learners of English, breaks down the study of passive voice into discrete phases, offering a thorough and engaging learning experience. In the first phase, students use the COCA Web Interface to access the Corpus of Contemporary American English (COCA) and carefully review real-world examples and student-generated phrases. The purpose of this dual analysis is to spot frequent problems that arise when learners switch from their home languages to English passive structures as well as intricacies in the usage of the passive voice in English.

The second step deviates from the norm by introducing a captivating role-playing scenario. With the help of COCA tools and Sketch Engine, students create contextually meaningful passive voice sentences. In order to ensure that students actively utilize passive voice structures in real-world circumstances, the role-playing scenario entails producing dialogues and textual exchanges. This practical practice promotes iterative skill development and improves passive voice production skills by providing instant feedback.

The last stage includes a collaborative component that encourages group investigation of passive constructions in BAWE – British Academic Written English corpus. To gain a deeper understanding of the difficulties faced by second language learners when utilizing passive voice structures in English, each group concentrates on a particular linguistic aspect, such as verb agreement or agent inclusion. This cooperative phase fosters critical thinking abilities and peer-to-peer knowledge exchange among students.

## 6. Mastering English Articles: A Corpus-Based Instructional Approach for Precision in Academic Writing

### 6.1 Learning Problem

In academic writing, using English articles correctly is crucial because it guarantees accuracy and coherence in the ideas presented. In addition to improving grammatical accuracy, proper article usage follows accepted academic traditions, which raises the writer's reputation and professionalism.

However, when it comes to academic writing, English learners face a tough challenge in becoming proficient, especially with using articles correctly. Butler (2002) argues that despite being the most frequently used words in English, articles like "a(n)" and "the" exhibit surprisingly intricate usage patterns. The absence of one- to-one form and meaning links in the English article system contributes significantly to the complexity associated with their usage. This is related to the fact that some languages lack a grammatical distinction between definite and indefinite articles, which requires more attention from learners when coping with these distinctions in English.

Another factor that contributes to the challenge of mastering article usage in English may be the absence of an analogous article system in the native language of English learners who lack articles (L1 article-less). This obviously leads to difficulties and potential interference from their first language (L1) (Chrzaszcz & Jiang, 2014). Additionally, individuals whose first language (L1) lacks an article system or possesses one significantly different from English, as observed in Turkish, may find it more challenging to use English articles (Lee Amuzie & Spinner, 2012).

### 6.2 Suggested Task

Via a corpus-based approach, this task aims to develop a sophisticated knowledge of English article usage among Turkish learners. The British National Corpus (BNC) has been selected as the corpus under investigation, and AntConc, a corpus analysis program, is the main instrument employed.

First of all, students will be granted access to the BNC through the BNC Web Interface. This web-based platform acts as a gateway to several genres, creating a clear connection between classroom learning and real-world English usage.

Following the introduction, students are expected to conduct an in-depth contextual examination inside the BNC. A sophisticated comprehension of contextual appropriateness is fostered by focusing on the intricate interactions between definite and indefinite articles in many linguistic contexts. By utilizing the BNC, the students will have the chance to explore articles in specific genres such as academic writing within the corpus. By doing this, they can have an insight into how articles are used and how context influences article choice.

Afterwards, to further promote students' learning, they will be introduced to AntConc, a powerful tool for corpus studies. By doing this, participants pinpoint recurring themes or omitted articles in their own language creation, which advances our comprehension of unique difficulties.

Following their familiarization with AntConc, the teacher should assign the students to create concordance lines for particular article-related terms (such as "a," "an," and "the") found in the corpus. By doing this, students will be able to examine the context in which these articles are found and talk about the subtle differences in how they should be used.

Add to this, within the scope of this web tool, students can be assisted in recognizing collocations connected to articles. They can be asked to investigate the common ways that articles are combined with nouns, adjectives, or verbs. With the use of this practice, students can better comprehend the frequent word pairings with articles.

## 7. Enhancing Academic Writing through Corpus-Informed Vocabulary Development

### 7.1 Learning Problem

Gaining vocabulary knowledge is a deeper matter than it seems to be. Most of the time, the learners only have the time to either learn one singular meaning of a word, or just memorize as they progress through a reading text. Knowing a word by its definition is only the tip of the iceberg. A learner can become conversationally fluent and proficient in a short period of time, but in many cases, achieving academic proficiency would take longer. To be proficient in written English, the learner has to build a strong vocabulary. Strengthening vocabulary skills include conceptualizing words, understanding the differences between spoken and written language items, and making use of different sources of information. The most widely available source is dictionaries; however, using a dictionary might be inconvenient for learners for various reasons: Learners might have difficulty in searching for or finding words. They also might not be able to understand the given definition(s) or choose the definition that best fits the context. In addition, they might forget the items they have looked for, and might have problems using the word in a sentence.

## 7.2 Suggested Task

One way of developing vocabulary is to adapt a corpus, which can be newly built or ready-made, such as the BNC (British National Corpus) Online and COCA (the Corpus of Contemporary American English). A corpus highlights how native speakers of English use certain language forms, vocabulary items, and expressions in real situations and puts an end to the so-called necessity of relying on a native speaker's intuition to tell what is commonly or rarely used (Dazdarevic et al., 2015). Of course, an individual's learning depends on their interests, needs, and motivation. Here, a learner corpus would be beneficial with how it is built upon the learners' use of language. Learner corpora can help assess the level of language development of the learners, as well as create an opportunity to analyze discourse features of learner language.

The implementation of a concordance tool is an excellent start to creating a learner corpus, especially to observe collocation and grammar errors. Corpus tools can be used to search for word frequencies and contexts, which can reveal how the learners' use differs from or aligns with the typical use of language items. An analysis of how often words are used, how they go together, and their different forms can help measure how advanced learners' vocabulary skills are. The scale of this development can be referred to as "lexical sophistication", which is also determined by the word frequencies. Lexical sophistication can be increased by using low-frequency words, which are likely to be hard (for some learner levels), and more formal words. This affects the "register" of the spoken or the written language. Register-based measures are closely relevant to the connection between the context and the vocabulary through the use of corpora-based software, or AWL (Academic Word List).

Learner corpora are easy to create from essays that the learners bring to the classroom or learner-produced texts. An example would be gathering a few essays from randomly chosen students during class and asking them to choose a number of words from each other's work, depending on the length of the essays. The words and the sentences that contain them are marked or written down before the students are asked to give the inflections (whenever possible), noun, adjective, or verb forms of the words. After all words are worked on in this fashion, the words and their typical uses are inspected via a corpus tool, such as COCA. Any missing forms, inflections, and collocations are added to the notes during this part. This part could be incredibly helpful in detecting errors in the concepts of collocations, grammar, and context. Expanding on the above example, the students are later asked to come up with words that could replace the ones that they choose; the words are searched in AWL or AVL for confirmation and further inspiration. The AWL can be created using a corpus (such as BNC Baby, BAWE, BASE, EAP, and ESP), or found readily available on the internet. The works are saved to be inspected after a few weeks to reveal the changes in the students' skills for both the students and the instructor(s) to observe. In this example, the instructor(s) can have a chance to examine the learners' discourse features, encourage collaborative learning by assigning groups while using the corpus tool, correct errors, track the learners' vocabulary development, and work on the register, which is a crucial element of developing academic skills.

## 8. Fostering Academic Formality: A Corpus Analysis Approach to Enhancing the Style

### 8.1 Learning Problem

Formality is a highly crucial part of academic writing which is closely related to word choice. It is of utmost importance to use formal equivalents of the colloquial words in an academic text. For the purpose of familiarizing the students with formal expressions, the corpus can be utilized. Through the use of activities that include the use of corpus, the teachers can help the students be aware of the formal academic expressions that they can use as a replacement for the words that are less suitable in an academic text. A corpus-based task that can be used to increase the students' awareness of their word choices in academic writing to enhance formality is presented below.

### 8.2 Suggested Task

As a first step of the task, the students are presented with a list of colloquial words that are compatible with their language proficiency level. After a list is provided to the students, the students are given a corpus consisting of sentences that include the formal words and phrases which can be used as a replacement for the colloquial words in the list. The relevant sentences can be gathered through the use of COCA, a freely available corpus of English, as it allows the searches to be limited to an academic genre. The students are then asked to find the formal equivalents of each colloquial word in the corpus data provided. At this stage of the task, the students can collaborate with each other to exchange their ideas. Upon the completion of the analysis of the texts, the students can be asked to use the formal expressions that they encounter in the corpus data in an academic context.

By doing so, the students can gain a better understanding of the use of correct academic words and expressions to enhance formality in their writing. As the last step of the task, the students can discuss the sentences that they produce using the corpus data with their peers and their teacher. This activity is in line with indirect instruction as the students are required to observe, analyze, and draw inferences from the corpus data. Such an activity allows them to use their creativity, which facilitates learning.

## 9. Conclusion

Focusing on three key areas being grammatical structural organization, formal language use, and appropriate word choice, the present paper has presented useful and effective corpus-based teaching tasks for teachers, bringing pedagogy into line with the changing requirements of academic writing. By utilizing real-world examples and practical exercises, students are able to improve their passive voice production skills and gain a deeper understanding of the intricacies of the English language. The collaborative component of this task also encourages peer-to-peer knowledge exchange and critical thinking, making it an effective and well-rounded learning experience. The corpus-based approach using the BNC and AntConc offers a comprehensive and dynamic way of learning English article usage for Turkish learners. By examining real-world English usage in various genres and utilizing powerful corpus analysis tools, students can gain a sophisticated understanding of how articles are used in different contexts. Moreover, the practice of creating concordance lines and recognizing collocations can help students to better comprehend the nuances of English article usage.

Overall, this method can be highly effective in fostering students' English language proficiency. Adapting a corpus and implementing a concordance tool is an excellent way to develop vocabulary, especially for language learners. Learner corpora are easy to create and can help assess the level of language development of the learners, as well as create an opportunity to analyze discourse features of learner language. By using corpus-based software, learners can improve their lexical sophistication, which is crucial for developing academic skills. The use of corpora-based software and AWL can help learners understand the connection between the context and the vocabulary. In general, the use of corpora-based software can be highly beneficial for language learners and instructors alike. Academic writing skills can also be improved by exposing students to formal words and phrases commonly used in academic settings. By analyzing the corpus data and using formal expressions in an academic context, students can gain a better understanding of how to enhance the formality of their writing.

Additionally, the collaborative aspect of the task encourages students to exchange ideas and work together, which can further enhance their learning experience. To put it another way, this activity is a valuable tool for teachers to use in the classroom to help students improve their writing skills and prepare them for academic success. In conclusion, this paper focuses on enhancing students' academic writing skills by developing a profound linguistic ability. The paper may provide teachers with effective tools and help students grasp the complexities of academic writing through the use of corpus-based tasks.

## References

To'laganova, A. Q. qizi, & Muratovna Panabayeva, M. (2022). *The Importance of Vocabulary in Language Learning and How to be Taught*. Central Asian Research Journal for Interdisciplinary Studies, 2(11). <https://doi.org/10.24412/2181-2454-2022-11-220-226>

Baleghizadeh, & Gordani. (2012, November 6). Academic Writing and Grammatical Accuracy The Role of Corrective Feedback. *Gist Education and Learning Research Journal*, 6(1692-5777), 159–176.

Brooks, G., Clenton, J., & Fraser, S. (2021). Exploring the importance of vocabulary for English as an additional language learners' reading comprehension. *Studies in Second Language Learning and Teaching*, 11(3), 351–376. <https://doi.org/10.14746/ssllt.2021.11.3.3>

Butler, Y. (2002). Second Language Learners' Theories On The Use Of English Articles: An Analysis of the Metalinguistic Knowledge Used by Japanese Students in Acquiring the English Article System. *Studies in Second Language Acquisition*, 24(3), 451-480. doi:10.1017/S0272263102003042

Cain, K., & Oakhill, J. (2014). *Reading comprehension and vocabulary: Is vocabulary more important for some aspects of comprehension?* L'Année Psychologique, 114(04), 647– 662. doi:10.4074/s0003503314004035

Chrabaścz, A., & Jiang, N. (2014, June 3). The role of the native language in the use of the English nongeneric definite article by L2 learners: A cross-linguistic comparison. *Second Language Research*, 30(3), 351– 379. <https://doi.org/10.1177/0267658313493432>

Dazdarevic, S., Zoranic, A.-L., & Fijuljanin, F. (2015). *Benefits of Corpus-Based Approach to Language Teaching*. Balkan Distance Education Network, (7). ISSN: 2334-9220 (Online), [https://www.researchgate.net/profile/SaminaDazdarevic/publication/282186652\\_BENEFITS\\_OF\\_CORPUSBASED\\_APPROACH\\_TO\\_LANGUAGE\\_TEACHING/links/56071e7908aea25fce3998ee/BENEFITS-OF-CORPUS-BASED-APPROACH-TO-LANGUAGE-TEACHING.pdf](https://www.researchgate.net/profile/SaminaDazdarevic/publication/282186652_BENEFITS_OF_CORPUSBASED_APPROACH_TO_LANGUAGE_TEACHING/links/56071e7908aea25fce3998ee/BENEFITS-OF-CORPUS-BASED-APPROACH-TO-LANGUAGE-TEACHING.pdf)

Halik, & Kareema. (2020, December). Difficulties faced by the ESL Learners in using English Passive Voice in Written Communication: A study based on the Students of NAITA, Trincomalee, Sri Lanka. *KALAM – International Research Journal Faculty of Arts and Culture*, 13(3)(27382214), 51–59. <http://ir.lib.seu.ac.lk/handle/123456789/5285>

Heylighen, F., & Dewaele, J.-M. (1999). *Formality of Language: Definition, measurement and behavioral determinants*. Internal Report, Center "Leo Apostel". Free University of Brussels, Belgium.

Hinkel, E., & Fotos, S. (Eds.). (2001). *New Perspectives on Grammar Teaching in Second Language Classrooms* (1st ed.). Routledge. <https://doi.org/10.4324/9781410605030>

Larsson, T., Kaatari, H., Dixon, T. & Egbert, J. (2023). Examining novice writers' perceptions of formality. *Journal of English for Research Publication Purposes*, 4(1), 29-55.

Lee Amuzie, G., & Spinner, P. (2012, December 7). Korean EFL Learners' Indefinite Article Use with Four Types of Abstract Nouns. *Applied Linguistics*, 34(4), 415–434. <https://doi.org/10.1093/applin/ams065>

Liardét, C. L., Black, S., & Bardetta, V. S. (2019) Defining formality: Adapting to the abstract demands of academic discourse. *Journal of English for Academic Purposes*, 381, 146–158

## Appendix 1: Abbreviations

BNC: British National Corpus

COCA: Corpus of Contemporary American English

AWL: Academic Word List

AVL: Academic Vocabulary List

BAWE: British Academic Written English Corpus

BASE: British Academic Spoken English Corpus

EAP: English for Academic Purposes

ESP: English for Specific Purposes