

**Gaukhar Yersultanova<sup>1\*</sup>, Guldana Seidaliyeva<sup>2</sup>, Ulbolsyn Kaiyrbekova<sup>3</sup>**<sup>1\*</sup>Corresponding author, Doctor of Philosophy (PhD), postdoctoral fellow,  
Abai Kazakh National Pedagogical University, Kazakhstan, Almaty,  
ORCID: 0000-0001-8913-6067 E-mail: gyersultanova@gmail.com<sup>2</sup>Doctor of Philosophy (PhD), Abai Kazakh National Pedagogical University,  
Kazakhstan, Almaty, ORCID: 0000-0002-4448-2414 E-mail: gseidaliyeva@gmail.com,<sup>3</sup>Doctor of Philosophy (PhD), Docent, People's Friendship University named after Academician  
A. Kuatbekov, Kazakhstan, Shymkent, ORCID: 0000-0002-9466-024X E-mail: ulbolsyn\_kudn@mail.ru

## APPLYING CORPUS-BASED METHODS TO ENHANCE ENGLISH LANGUAGE LEARNING

**Abstract.** This study aimed to investigate the effectiveness of Corpus of Contemporary American English (COCA) corpus-based activities on the performance of Kazakhstani students in English Language Teaching (ELT) and the learners' attitudes regarding the corpus-based approach. Thirty-three 2<sup>nd</sup> year students at Al-Farabi Kazakh National University participated in the study (proficiency level in English was Intermediate (B1)). The second-year students were randomized into two intact groups, one for experimentation and the other for control. The participants were expected to learn three grammar structures and vocabulary related to technology, podcasts, and art by the end of the study. Participants in the control group learned grammar and vocabulary structures through coursebook materials, whereas those in the experimental group learned these concepts through corpus-based exercises. The data were collected through pre-test, post-test, and an 8-item-questionnaire based on a 5-point Likert scale as quantitative instruments. According to a statistical analysis of the post-test findings, the corpus-based exercises proved more successful than the coursebook contents. Furthermore, the approach questionnaire analysis showed that students' views regarding applying corpus-based learning activities in ELT were neutral.

**Keywords:** corpus-based activities; English language teaching; vocabulary; grammar structures; questionnaire**For citation:** Yersultanova, G., Seidaliyeva, G., Kaiyrbekova, U. Applying Corpus-Based Methods to Enhance English Language Learning. *Tiltanyim*, 2024. №4 (96). P. 164-176.DOI: <https://doi.org/10.55491/2411-6076-2024-4-164-176>**Гаухар Ерсултанова<sup>1\*</sup>, Гулдана Сейдалиева<sup>2</sup>, Ұлболсын Қайырбекова<sup>3</sup>**<sup>1\*</sup>автор-корреспондент, философия докторы (PhD), постдокторант,  
Абай атындағы Қазақ ұлттық педагогикалық университеті, Қазақстан, Алматы қ.,  
ORCID: 0000-0001-8913-6067 E-mail: gyersultanova@gmail.com<sup>2</sup>философия докторы (PhD), Абай атындағы Қазақ ұлттық педагогикалық университеті,  
Қазақстан, Алматы қ., ORCID: 0000-0002-4448-2414 E-mail: gseidaliyeva@gmail.com<sup>3</sup>философия докторы (PhD), доцент, Академик А. Қуатбаев атындағы  
Халықтар достығы университеті, Қазақстан, Шымкент қ.,  
ORCID: 0000-0002-9466-024X E-mail: ulbolsyn\_kudn@mail.ru

## АҒЫЛШЫН ТІЛІН ҮЙРЕНУДІ ЖАҚСARTУ ҮШІН КОРПУСТЫҚ ӘДІСТІ ҚОЛДАНУ

**Андатпа.** Зерттеудің негізгі мақсаты – қазақстандық студенттерге Қазіргі американдық ағылшын тілі корпусы (COCA) материалдарын қолдану арқылы ағылшын тілін оқыту және оқушылардың корпустық әдісті қолдану барысындағы көзқарастарын зерттеу. Бұл зерттеуге Әл-Фараби атындағы Қазақ ұлттық университетінде оқитын отыз үш екінші курс студенті қатысты (ағылшын тілін меңгеру деңгейі – B1). Екінші курс студенттері кездейсоқ түрде екі топқа бөлінді: тәжірибелік және бақылау топтары. Зерттеу нәтижесінде студенттер үш грамматикалық құрылым мен технология, өнер мен подкастарға қатысты жаңа вокабулярды меңгереді деп күтілді. Тәжірибелік топтағы студенттер аталып өткен грамматикалық құрылымдар мен сөздікті корпустық материалдарды қолдану арқылы үйренді; ал бақылау тобы тек стандартты оқулықтарды ғана қолданды. Деректерді жинау және сандық талдау алдын ала тестілеу, кейінгі тестілеу және Лайкерттің бес балдық шкаласы бойынша бағаланған сегіз сұрақтан тұратын сауалнама арқылы жүзеге асырылды. Тестілеуден кейінгі статистика стандартты оқулықтардағы жаттығулармен салыстырғанда корпус материалдарын пайдаланудың тиімділігін анықтады. Сонымен қатар сауалнама нәтижелерін талдау деректері студенттердің ағылшын тілін оқытуда тілдік корпустарды қолдануға бейтарап көзқарасын анықтады.

**Түйін сөздер:** корпуслық жаттығулар; ағылшын тілін оқыту; сөздік қор; грамматикалық құрылымдар; сауалнама

**Сілтеме жасау үшін:** Ерсұлтанова Г., Сейдалиева Г., Кайырбекова Ұ. Ағылшын тілін үйренуді жақсарту үшін корпуслық әдісті қолдану. *Tiltanyum*, 2024. №4 (96). 164-176-бб. (ағыл. тілінде)

DOI: <https://doi.org/10.55491/2411-6076-2024-4-164-176>

**Гаухар Ерсұлтанова<sup>1\*</sup>, Гулдана Сейдалиева<sup>2</sup>, Улболсын Кайырбекова<sup>3</sup>**

<sup>1\*</sup>автор-корреспондент, доктор философии (PhD), постдокторант,

Казахский национальный педагогический университет имени Абая, Казахстан, г. Алматы,

ORCID: 0000-0001-8913-6067 E-mail: gyersultanova@gmail.com

<sup>2</sup>доктор философии (PhD), Казахский национальный педагогический университет имени Абая,

Казахстан, г. Алматы, ORCID: 0000-0002-4448-2414 E-mail: gseidalieva@gmail.com,

<sup>3</sup>доктор философии (PhD), доцент, Университет дружбы народов имени

академика А. Куатбекова, Казахстан, г. Шымкент,

ORCID: 0000-0002-9466-024X E-mail: ulbolsyn\_kudn@mail.ru

## ПРИМЕНЕНИЕ КОРПУСНЫХ МЕТОДОВ ДЛЯ УЛУЧШЕНИЯ ИЗУЧЕНИЯ АНГЛИЙСКОГО ЯЗЫКА

**Аннотация.** Данное исследование проведено с целью выявления эффективности использования корпусных материалов в обучении казахстанских студентов английскому языку на основе Корпуса современного американского английского языка (COCA), а также изучения отношения обучающихся к использованию данного метода при обучении английскому языку. В данном исследовании приняли участие тридцать три студента 2 курса Казахского национального университета имени Аль-Фараби (уровень владения английским языком – B1). Студенты 2 курса были разделены методом случайной выборки на две группы: экспериментальная и контрольная. По окончании эксперимента участники должны были усвоить 3 грамматические структуры и новый вокабуляр, связанный с технологией, искусством и различными подкастами. Участники экспериментальной группы изучали грамматику и словарь с использованием упражнений, основанных на материалах корпуса, в то время как контрольная группа продолжала изучать грамматику и лексический минимум с использованием только материалов из стандартных учебников. Сбор и количественный анализ данных осуществлялся с помощью предварительного тестирования, пост-тестирования, и опросника, состоящего из восьми вопросов, который оценивался по пятибалльной шкале Лайкерта. Статистические данные пост-тестирования выявили эффективность использования корпусных материалов по сравнению с упражнениями из стандартных учебников. Кроме этого, данные анализа результатов опроса выявили нейтральное отношение студентов к использованию языковых корпусов в обучении английского языка.

**Ключевые слова:** корпусные упражнения; преподавание английского языка; вокабуляр; грамматические структуры; опросник

**Для цитирования:** Ерсұлтанова Г., Сейдалиева Г., Кайырбекова Ұ. Применение корпусных методов для улучшения изучения английского языка. *Tiltanyum*, 2024. №4 (96). С. 164-176. (на англ. яз.)

DOI: <https://doi.org/10.55491/2411-6076-2024-4-164-176>

## Introduction

Corpus linguistics principles are increasingly used in English language learning and teaching, with scholars recognizing them as essential tools for enhancing teaching effectiveness and fostering active discourse in classroom settings (Hunston, 2002). As a result, corpora are integrated into classroom practices in two primary ways: directly, through engaging activities that are designed around the corpus itself, and indirectly, through materials and products such as coursebooks and dictionaries that are informed by corpus data.

The development of corpus-based resources significantly transformed language education, particularly with the launch of the Collins COBUILD English Language Dictionary in 1987, which set a new standard for integrating authentic language data into learning materials (Hunston, 2002; O’Keeffe et al., 2007). This milestone underscored the growing relevance of corpora in teaching, fostering an increasing interest in their application among educators and researchers. The Teaching and Language Corpora (TaLC) conference has since emerged as a vital forum for sharing innovative research and practices related to corpus-based language education. By promoting collaboration among linguists and educators, the conference has played a key role in advancing data-driven learning (DDL), a methodology that leverages corpus data to enhance language acquisition (Bernardini, 2004; Boulton,

Cobb, 2017). Through these efforts, corpus-based approaches have become integral to modern language teaching, bridging the gap between theory and practice.

Despite an extensive body of research examining the implications of corpus-based exercises and their effects on learners' performance within English Language Teaching (ELT), a significant gap persists regarding students' attitudes toward engagement in corpus-based activities. This study aims to address this deficiency by investigating the impact of corpus-based exercises on the performance of Intermediate (B1) English as a Foreign Language (EFL) students in ELT contexts. Furthermore, the research seeks to assess learners' perceptions and attitudes toward this innovative pedagogical approach, thereby providing valuable insights into their experiences and the potential for future applications within language education.

Historically, language acquisition has hinged on traditional pedagogical frameworks emphasizing textbooks and instructor-generated materials, which typically prioritize theoretical constructs over practical application. However, in the context of our increasingly information-saturated and digitized environment, corpus-based language teaching (CBLT) has emerged as a progressive alternative. CBLT leverages extensive corpora-comprehensive datasets of authentic language use encompassing written texts, spoken dialogues, and diverse communicative forms. These corpora afford learners critical insights into the pragmatic functioning of language in real-world settings. By prioritizing authentic language examples over prescriptive grammar and vocabulary lists, CBLT facilitates a deeper engagement with linguistic nuances, idiomatic expressions, and the contextual subtleties of meaning that characterize everyday communication. This methodology operates not merely as a replacement for traditional approaches but serves as a valuable adjunct, enriching existing instructional practices.

Integrating corpora into the curriculum allows educators to enhance their pedagogical techniques while fostering critical analytical skills in learners. Students learn to interpret language within practical scenarios, thereby cultivating their linguistic intuition. The use of authentic data enhances the relevance and efficacy of language instruction, as it enables learners to discern usage patterns, observe contextual variations, and familiarize themselves with diverse dialects and registers. As technological advancements lower the barriers to collecting and analyzing large datasets, the potential for CBLT continues to expand. The implications of this instructional paradigm are profound; it holds the promise of transforming language education while equipping learners to thrive in a diverse and interconnected global landscape, where comprehending authentic language in contextualized settings is paramount for effective communication. In this rapidly evolving educational milieu, CBLT exemplifies the necessity of adapting teaching methodologies to address the complexities of modern language use and the needs of contemporary learners.

The research study is focused on addressing the following questions:

1. What is the impact of employing corpus-based activities on the academic achievement of Kazakhstani students in English Language Teaching (ELT)?
2. What are the perspectives of Kazakhstani students regarding the corpus-based approach in ELT?

## **Literature review**

### *Corpora and Language Teaching*

Language instructors and students have increasingly embraced corpus-based approaches in English Language Teaching (ELT), utilizing tools such as corpus-informed dictionaries and grammar exercises to enhance language learning. A notable catalyst for this development is the biennial TaLC (Teaching and Language Corpora) conference, which has provided a platform for linguists and educators to explore the pedagogical potential of collocations and concordances. One significant contribution to this field is the textbook *Concordances in the Classroom* by Tribble and Jones (1990), which emerged from early TaLC discussions and remains a foundational resource.

Corpus Linguistics (CL) serves as a key source of data, methodologies, and insights for language teaching, underscoring the symbiotic relationship between CL and language instruction. Leech delineates two principal categories of corpus use in language teaching: direct and indirect applications (Leech, Candlin, 1997). Direct applications refer to the explicit use of corpora by educators and learners for instructional purposes (Bernardini, 2002), while indirect applications involve the incorporation of

corpus-derived materials, such as textbooks and teaching resources, into the classroom (Hunston, 2002). Figure 1 illustrates how these direct and indirect approaches can be effectively integrated into language teaching practices, providing educators with a range of strategies to enrich their pedagogical techniques.

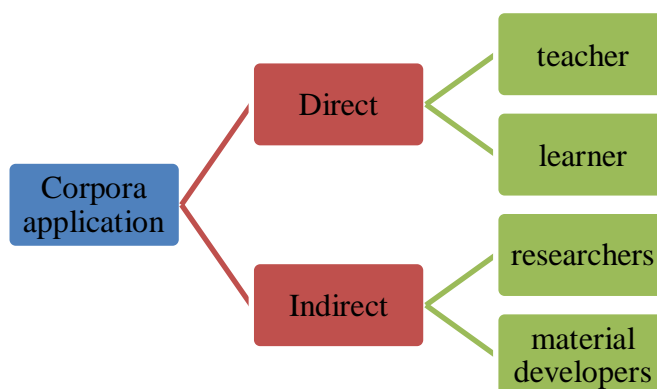


Figure 1 – Application of corpora in language teaching  
 Сурет 1 – Тілді оқытуда корпусты қолдану  
 Рисунок 1 – Применение корпусов в преподавании языка

*Utilizing a corpus-based approach for vocabulary instruction.*

The corpus-based approach to vocabulary learning has garnered significant attention in language education for its ability to enhance learners' engagement and comprehension of language in context. By interacting with authentic language data, students benefit from exposure to real-world examples, promoting deeper understanding, vocabulary acquisition, and retention.

*Benefits of Corpus-Based Vocabulary Learning*

One of the key advantages of the corpus-based approach is its impact on vocabulary acquisition. Studies indicate that learners utilizing corpus-based materials show considerable improvements in vocabulary knowledge compared to those relying on traditional methods (Alenizi, Adawi, 2024; Tosun, Sofu, 2023). This improvement is attributed to the exposure to authentic contexts, which facilitates a more nuanced grasp of word meanings and usage. Furthermore, advanced learners particularly benefit from corpus approaches in recognizing and producing academic vocabulary, leading to the development of lexical competence (Can, 2023).

Learner attitudes towards the corpus-based approach are generally positive. Many students appreciate the opportunity to interact with authentic language samples, finding value in observing how vocabulary functions in real-world settings (Alenizi, Adawi, 2023). This engagement not only fosters a favorable learning experience but also motivates learners to independently explore language patterns.

*Limitations of Corpus-Based Vocabulary Learning*

Despite its benefits, the corpus-based approach is not without challenges. Data overload poses a significant issue, as the vast amount of information in corpora can overwhelm learners, especially those with limited proficiency (Tosun, Sofu, 2023). Without proper guidance, learners may struggle to extract relevant data, impeding their learning process.

Another limitation is the necessity for specialized training in using corpus tools effectively. Alenizi and Adawi (2024) note that a lack of training resources in some EFL (English as a Foreign Language) contexts limits the practical implementation of corpus-based learning. This challenge is particularly pronounced in resource-constrained settings where instructors may lack the expertise or tools to guide learners.

Moreover, while corpus-based approaches often yield improvements in vocabulary acquisition, some studies report mixed results regarding their statistical significance compared to conventional methods (Can, 2023). These findings suggest that the effectiveness of corpus-based learning depends on factors such as the learners' proficiency levels and the instructional design.

The study by Aied Alenizi and Reem Adawi (2024) investigated the effectiveness of corpus-based materials in enhancing vocabulary learning among Saudi EFL students. The research involved 24



female participants, who were divided into experimental and control groups for a five-week period. The experimental group used corpus-based resources, specifically the COCA corpus, while the control group followed traditional textbook-based instruction. The results showed that the experimental group significantly outperformed the control group in vocabulary acquisition, as demonstrated by higher post-test scores. Furthermore, a questionnaire revealed that most students had positive attitudes toward using corpus-based materials, despite some initial difficulties. These findings underscore the potential of corpus-driven approaches to improve vocabulary retention and create a more engaging learning experience. The study also calls for further research to study the broader applicability of corpus-based learning in different educational contexts.

In a similar vein, Youssef (2020) investigated the effectiveness of a corpus-based approach in enhancing vocabulary learning and retention among tertiary-level EFL learners in Saudi Arabia. The study utilized tailored corpus-based materials and activities designed to promote vocabulary acquisition, focusing on authentic language use and word patterns. The findings revealed significant vocabulary learning gains, as evidenced by improvements in pre- and post-test scores. Moreover, the corpus-based tasks not only supported learners in mastering discipline-specific terminology but also fostered long-term retention of vocabulary items. Participants also expressed positive feedback, appreciating the exposure to real-world language examples and the relevance of the materials to their academic and professional needs.

Both studies underscore the effectiveness of corpus-based approaches in enhancing vocabulary acquisition. They demonstrate that utilizing concordance lines and specialized corpora provides learners with meaningful, authentic input, which not only improves vocabulary retention but also deepens their understanding of language use. Moreover, learners' positive attitudes, as captured through Likert scale surveys and interviews, highlight the perceived value of these methods in supporting their academic and professional goals. These findings reinforce the potential of corpus-based approaches to make vocabulary learning more engaging and impactful, particularly when tailored to learners' specific needs and contexts.

#### *Corpus-based methodology for grammar instruction*

The importance of teaching grammar lies in its ability to structure communication, connecting words and sentences to convey meaning effectively. Corpus-based approaches in grammar instruction offer opportunities to explore collocations and grammatical patterns in authentic contexts, thereby enhancing the meaningfulness of language input. According to Sun and Wang (2020), integrating corpus tools in classrooms fosters a deeper understanding of grammar and vocabulary by presenting students with real-life examples of language use.

A study by Liu and Jiang (2009) evaluated the use of corpus-based approaches in teaching lexicogrammar, involving 244 EFL participants across Chinese and U.S. universities. They employed concordance lines and collocations as teaching materials. The post-study questionnaire revealed that 59.6% of participants found corpora helpful for understanding word meanings and structural usage patterns, while 69.7% noted improved ability to distinguish synonyms. However, challenges included the overwhelming volume of corpus data and irrelevant examples for specific fields of study. To address these issues, Liu and Jiang emphasized the importance of curating relevant concordance lines and incorporating group work to enhance engagement in large classes.

In another study, Girgin (2011) explored corpus-based grammar teaching at Erciyes University, focusing on EFL learners with lower proficiency levels. The experimental groups replaced coursebook activities with tasks derived from the Corpus of Contemporary American English (COCA). While pre- and post-tests revealed similar knowledge gains between experimental and control groups, learners' attitudes toward corpus-based instruction were neutral. Girgin highlighted the necessity of preparing tailored materials for learners with limited proficiency to maximize the approach's effectiveness.

Both studies illustrate the potential of corpus-based activities to enhance grammar instruction, with notable benefits such as improved synonym differentiation and structural comprehension. However, effective implementation requires addressing challenges like data overload and ensuring the relevance of materials to learners' needs. These findings suggest that while corpus-based grammar instruction is promising, its success depends on careful design and integration into the curriculum (Liu,

Jiang, 2009; Girgin, 2011; Su, Wang, 2020).

## **Materials and methods**

### *Participants*

The participants of this experiment were thirty-three 2<sup>nd</sup> year students of Al-Farabi Kazakh National University. The age range of the participants was 17 to 20. There were 16 female students and 17 male students in the study. The average level of proficiency in English of the participants was Intermediate (B1).

The study was conducted during the fall term of the academic year 2023-2024. The researcher finished the curriculum for two complete courses as part of the study's goal to assess the efficacy of corpus-based activities. The entire classes were divided randomly into two groups for the study: the experimental group, comprising seventeen students, and the control group, comprising sixteen students.

### *Instruments and Materials*

The experimental study utilized pre-test, post-test, and attitude questionnaire instruments, which were derived from previous research. The majority of these instruments were adapted from prior studies and were adjusted to align with the current research objectives. The study utilized supplementary materials including corpus-based activities, a coursebook, and COCA, in addition to the instruments. The Head of the Department of Foreign Philology and Translation Studies at al-Farabi KazNU approved all the supplementary materials.

### *Pre-test and Post-test*

The pre-and post-tests were conducted to assess the effectiveness of the corpus-based method in addressing the initial study question. Three target grammar structures and vocabulary structures were selected from the syllabus to create questions for before and after a test. These grammar structures were the passive voice, reported speech and the conditional sentences.

Regarding the vocabulary, new words related to AI and modern technology in education, podcasts and TED talks, and art were selected. The questions on the pre-tests and post-tests were of the "filling the gap" type.

The pre-tests and post-tests comprised 10 questions and were assessed online using Google Forms. Each question was worth 1 point, except for the third conditional question, which allowed for partial scores (i.e., 0.5 and 1) for partially correct answers.

### *An Attitude Questionnaire*

The students in the experimental group were instructed to fill out an attitude questionnaire to address the second research question. Since respondents' attitudes are frequently gauged using the Likert scale, the questionnaire was formatted in this manner. In addition, participants can select the Likert scale option that most closely matches their perspective. The questionnaire's items were modified from Girgin's (2011) study. However, the 6-point scale was modified to 5-point scale and 11 questions were compressed into 8 questions. Table 1 below outlines the focus of the questionnaire items and the 5-point scale, where 1 point describes a strongly negative attitude, 3 points describe neutrality, and 5 points describe strongly positive attitudes. According to the items, Item 1 and Item 2 are rating the corpus-based activities directly, through difficulty level and usefulness. Following that, Item 3 and Item 4 are comparing corpus-based activities to a coursebook, in the frame of difficulty and boredom. Other items (i.e., Item 5, Item 6, Item 7, Item 8) are more personal, where students rate their achievements, compare preferences, add their own attitudes and recommendations towards corpus-based activities.

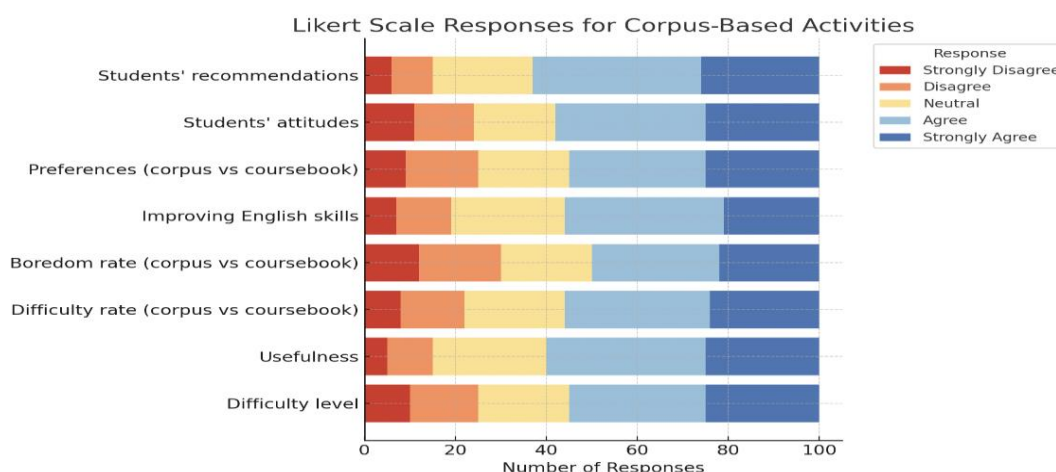


Figure 2 – The focus of questionnaire items  
 Сурет 2 – Сауалнама элементтерінің фокусы  
 Рисунок 2 – Основное внимание в опросниках

### Materials

To teach target grammar and vocabulary structures, instructors typically use coursebooks. After that, the control group's students were instructed to use the coursebook, while the experimental group was instructed not to use it.

In the experimental group, traditional assignments were substituted with corpus-based activities, allowing students to engage directly with the COCA. The concordance lines for grammar structures were adapted from Girgin's (2011) study and the researcher created the vocabulary structures. Concordance lines selected for this study were used for presentation and practice. It is evident from the data depicted in the figure 3, that concordance lines provide students with meaningful input, that students guess, analyze, and synthesize the meanings of the new structures. In the corpus-based approach, the role of the teacher is a facilitator; the teacher usually facilitates discussion for a broad understanding of target structures. Furthermore, the teacher guides students to the appropriate discussions. For instance, the figure below demonstrates lines followed by leading questions for deeper understanding.

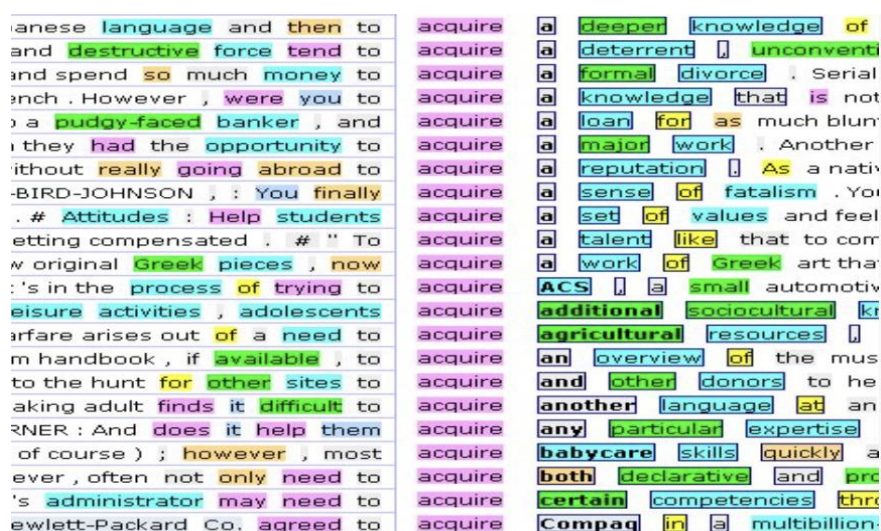


Figure 3 – Concordance lines for teaching vocabulary  
 Сурет 3 – Сөздік қорды оқытуға арналған сәйкестік сызықтары  
 Рисунок 3 – Линии соответствия для обучения лексике

### *Procedure*

After obtaining the required signatures on the consent letters and fulfilling the required authorization processes with the English Department at KazNU, three target grammar structures and target vocabulary structures were selected according to the syllabus, to prepare the materials and instructions for the learners.

In the first stage of the research, the researcher sent the pre-test questions online using Google Forms. Overall, filling the gaps of the pre-test took roughly 10 minutes. Google Forms were used for a study to teach grammar structures and vocabulary to students. The experimental group was given a demo lesson explaining the corpus approach and COCA. Concordance lines were prepared beforehand due to lack of laptops. The experimental group was not allowed to look at books or share corpus-based activities, while the control group used a coursebook and prepared materials. A post-test was conducted using Google Forms for both groups, with the same difficulty levels. The experimental group's attitude questionnaire was used to share their attitudes and recommendations towards the approach, as only experimental participants could evaluate it.

### *Data Analysis*

Quantitative analysis of the post-test results was used to address the study's first research question. To assess the progress of the experimental and control groups, the researcher relied on statistical analysis using a Microsoft Excel spreadsheet to evaluate the scores. To assess the overall acquisition of the target grammar and vocabulary structures and to demonstrate the effectiveness of the corpus-based activities, the pre-test and post-test scores of both the experimental and control groups were meticulously compared.

In order to answer the second research question, the researcher used quantitative analysis of an attitude questionnaire. The mean and standard deviation were measured to find out the attitudes towards the corpus-based approach.

### **Results and discussions**

#### *Examination of Pre-test and Post-test Outcomes*

In Table 1, the pre-test and post-test scores of the experimental and control groups are compared. The table includes the medians, means, interquartile ranges (IQR), and standard deviation (Std. Dev.) of the scores. The results demonstrate the effectiveness of utilizing corpus-based activities in English Language Teaching (ELT) within the experimental group.

Table 1 – Comparing the control group with the experimental group

Кесте 1 – Бақылау тобын экспериментальді тобымен салыстыру

Таблица 1 – Сравнение контрольной группы с экспериментальной группой

<i>Group</i>	<i>Pre-Test Median</i>	<i>Post-Test Median</i>	<i>Pre-Test Mean</i>	<i>Post-Test Mean</i>	<i>Pre-Test IQR</i>	<i>Post-Test IQR</i>	<i>Pre-Test Std. Dev.</i>	<i>Post-Test Std. Dev.</i>
Experimental Group	7	8	6.4	8	1	2	1.13	1.83
Control Group	6.5	7.5	6.5	6.5	2	1	1.34	3.45

In the beginning Table 1's medians seem to indicate that both groups improved by one point in the post-test results. This indicates that while the control group's scores for the post-tests were equal to 6.5 and 7.5, respectively, the pre-test results showed that a majority of students in the experimental group scored 7 points, while in the later session, they scored 8 points. Furthermore, the average number indicates the group average and provides additional context for the data analysis. The experimental group's mean number rose from 6.4 to 8, while the control group's mean number stayed at 6.5. Given that both groups performed comparably in the pre-test, the experimental group's higher mean number indicates progress.

Moreover, IQR and Std. Dev. scores are also considered as important, due to the spread of the results. The interquartile range (IQR) of the experimental groups expanded, indicating greater variability in results across individual students. In contrast, the interquartile range (IQR) of the control



groups decreased, leading to reduced variability in the post-test results compared to the pre-test results. Furthermore, Std. Dev. was calculated to see how test scores were close to the mean value. As a result, the experimental groups' standard deviation slightly increased from 1.13 to 1.83, in contrast, the control groups' standard deviation rocked from 1.34 to 3.45. The control group's posttest results were highly variable, with students scoring both the minimum and maximum scores.

#### *Examination of the Attitude Survey*

Table 2 below illustrates students' attitudes toward corpus-based activities in ELT, using the Likert scale format. Seven students of the experimental group responded to that questionnaire. The overall average of 3.09 indicates a neutral attitude toward the corpus-based approach. The following paragraphs contain a more detailed interpretation of the questionnaire results.

Table 2 – Frequencies for items 1-8 in the attitude questionnaire

Кесте 2 – Көзқарас сауалнамасындағы 1-8 тармақтары бойынша жиіліктер

Таблица 2 – Частоты пунктов 1-8 в анкете по отношению

Items	1	2	3	4	5	Mean	Std. Dev.
1. I think that learning English language through corpus-based activities is/ isn't difficult.	Very difficult ← 00 5 2 0 → Very easy					3.29	0.49
2. Learning English language through corpus-based activities is/isn't useful.	Very useless ← 0 2 1 2 2 → Very useful					3.57	1.27
3. I think that learning English language through corpus-based activities are more difficult than learning English through a course book.	Strongly disagree ← 0 2 1 4 0 → Strongly agree					3.29	0.95
4. I think that learning English language through corpus-based activities are more boring than learning English through a course book	Strongly disagree ← 1 1 3 2 0 → Strongly agree					2.86	1.07
5. Using corpus-based activities in learning of English language can improve my English skills.	Strongly disagree ← 1 0 2 3 1 → Strongly agree					3.43	1.27
6. I prefer using corpus-based activities in learning of English language to using a course book in learning of English language.	Strongly disagree ← 1 3 3 0 0 → Strongly agree					2.29	0.76
7. I would really feel positively towards using corpus-based activities in learning of English language.	Strongly disagree ← 0 3 2 2 0 → Strongly agree					2.86	0.91
8. I recommend that teachers should use corpus-based activities so as to teach grammar/vocabulary/writing in the classroom settings.	Strongly disagree ← 0 2 2 3 0 → Strongly agree					3.14	0.91
Total:						3.09	0.95

Item 1 and Item 2 above find out the learners' attitudes towards the difficulty level and usefulness of corpus-based activities respectively. The vast majority of the participants (24) have chosen a neutral attitude in Item 1, the other participants concurred that activities based on a body of language were easy to use and there were no negative responses from anyone. Regarding Item 2 overall mean (3.57) shows that the corpus-based activities were found to be beneficial by the learners. Approximately half of the respondents (17) believed that those activities were useful, even very useful, whereas six participants

considered those useless, and three participants' attitude was neutral.

Subsequently, the analysis in Item 3 and Item 4 evaluates learners' perceptions of the difficulty and boredom levels associated with corpus-based activities in comparison to coursebook activities. Generally, the mean (3.29) of Item 3 demonstrated a neutral attitude, however, Std. Dev. (0.95) shows that the twelve participants stated corpus-based activities as difficult in comparison to the coursebook materials, whereas six participants claimed those to be easy. Concerning Item 4 the mean (2.86) and Std.

Dev. (1.07) demonstrates a neutral attitude towards the boredom rate, where roughly half of the participants (17) responded neutrally, however, twelve participants stated that corpus-based activities as interesting, and other participants (8) stated those as boring in comparison to the coursebook materials.

Finally, regarding more personal questions, remaining items were addressed to the participants. Item 5 demonstrates the highest Std. Dev. (1.27), and shows contrasting opinions. On the one hand, according to the data, one participant expressed strong disagreement regarding the potential of corpus-based activities to enhance English language skills, and conversely, another participant expressed strong agreement on that particular issue. According to other participants, corpus-based activities were found to enhance their English language proficiency. As for Item 6, three participants disagreed to prefer corpus-based activities to coursebook materials, and even one participant strongly disagreed. That means learners did not prefer to replace coursebook materials with corpus-based activities. As a result, three participants of Item 7 showed negative attitudes, two participants displayed neutral attitudes while the others exhibited positive attitudes, however, the mean (2.86) permits us to view those as neutral attitudes. In the final Item 8 three participants recommended using those activities to the teachers, two participants did not and the other two participants neutrally responded to the questionnaire. The mean (3.14) allows considering the recommendations as neutral.

Overall, there are still half-neutral and opposing views that ought to be taken into consideration, even though the corpus-based activities demonstrated their effectiveness and half of the participants agreed that it can improve their English language skills. The rate of boredom with corpus-based exercises and the substitution of coursebook contents for corpus-based activities were the most frequently reported instances of profoundly negative views. The responses hold significance for the materials development segment, as the participants suggested that language educators employ this methodology.

### *Discussion*

The findings of this study offer valuable insights into the use of corpus-based activities in English Language Teaching (ELT) in Kazakhstan. The results both align with and diverge from earlier studies, providing a nuanced understanding of the impact of corpus-based approaches on student learning and attitudes.

### *Efficacy of Corpus-Based Activities*

The first research question, which explored the efficacy of corpus-based activities, yielded positive results, confirming previous studies that highlighted the effectiveness of such activities in improving English language skills (Girgin, 2011; Paker, Ozcan, 2017). The significant improvement in the post-test scores of the experimental group suggests that corpus-based activities can be a powerful tool in enhancing language proficiency, particularly in vocabulary acquisition and grammatical understanding. These results are consistent with the research of (Hou, 2014), who found that corpus-based approaches facilitate more autonomous learning and greater exposure to authentic language use, both of which are essential for language learners. The positive results from this study suggest that incorporating corpora into the curriculum can help students build more accurate and nuanced language skills, particularly in contexts where traditional coursebook materials may fall short in providing authentic, real-world language examples.

However, it is important to acknowledge that while the experimental group showed improvement, the control group also exhibited progress, which could be attributed to other factors, such as the general pedagogical methods used or the students' prior knowledge of English. Future research could explore additional variables that might contribute to language improvement, such as the role of teacher guidance, peer interaction, or the use of technology in enhancing corpus-based activities.

### *Attitudes Toward Corpus-Based Activities*

The second research question, which assessed students' attitudes toward corpus-based activities, revealed a more complex picture. While previous studies with undergraduate and preparatory students (Girgin, 2011; Paker, Ozcan, 2017) reported predominantly positive reactions, the results of this study show that Kazakhstani students displayed a more neutral attitude toward corpus-based activities. The Likert scale data indicated that, while some students were receptive to these activities, a significant portion of the students expressed skepticism, especially when contrasting corpus-based activities with traditional coursebook materials.

This finding suggests that students may still be more accustomed to the structured and familiar approach of coursebooks, which are often easier to navigate and less demanding in terms of independent analysis. The negative attitudes expressed in Table 2 further highlight that some students may feel overwhelmed by the more self-directed nature of corpus-based activities, which require a higher level of cognitive engagement and critical thinking. This could be indicative of a broader challenge in implementing corpus-based activities in environments where students are more accustomed to teacher-directed, content-heavy lessons.

Interestingly, despite the neutral or negative reactions to the activities, students still recommended their inclusion in future lessons. This paradoxical finding suggests that, while students may not fully embrace corpus-based activities, they can see their potential value and recognize the benefits of exposure to authentic language data. This indicates that the challenge lies not in the potential of corpus-based activities themselves but in how they are integrated into the curriculum and how students' attitudes toward these activities are shaped.

### *Implications for Teaching Practice*

The mixed results regarding student attitudes suggest several practical implications for ELT practice in Kazakhstan. First, teachers may need to provide more scaffolding when introducing corpus-based activities, helping students understand their purpose and how to effectively use corpora to improve their language skills. A gradual transition from traditional coursebook materials to corpus-based tasks, with clear guidance and support, may help reduce students' resistance and increase their confidence in using corpus tools.

Furthermore, integrating corpus-based activities in a way that complements coursebook materials rather than replacing them entirely may be a more effective approach. This hybrid method could address students' preferences for familiar learning tools while still providing the benefits of authentic language input. In this context, teachers could use corpora to supplement coursebook exercises, providing students with real-world examples that reinforce language points presented in the textbooks.

### *Limitations and Directions for Future Research*

This study has several limitations that warrant further investigation. While students from Kazakhstan demonstrated proficiency in using corpus-based activities to enhance their linguistic abilities, several challenges were identified. First, larger sample sizes and additional groups are needed for more robust data analysis. Second, limited access to electronic devices in some university settings may hinder the integration of corpus-based methods. Third, these activities are better suited for self-directed learners, emphasizing the need for user-friendly materials developed by stakeholders. Fourth, extended teaching time is required to cover grammar and vocabulary structures thoroughly for more valid results. Additionally, qualitative data, such as interviews or open-ended surveys with students and teachers, could offer deeper insights into perceptions and experiences with corpus-based activities.

Furthermore, the study primarily examined the short-term effects of corpus-based activities. Future research should investigate their long-term impact on language development, exploring how students' attitudes and proficiency levels evolve with sustained exposure to such approaches. This would provide valuable insights into the efficacy and adaptability of corpus-based methods in ELT.

### *Pedagogical Implication*

The research emphasizes the efficacy of incorporating corpus-based activities in educational instruction, citing their value as a natural language source and increased meaningful input for learners. It recommends including such activities in coursebooks or preparing them as supplementary material. Independent learners can contribute to the field by completing projects related to corpus linguistics and

generating new findings.

The utilization of a corpus-based approach in English language teaching represents a significant shift in pedagogical practices. By systematically analyzing authentic language use drawn from large corpora, educators can more effectively understand the linguistic phenomena that occur in real-world communication. This approach facilitates a deeper engagement with language, allowing learners to encounter diverse lexical choices, grammatical structures, and contextual usage patterns.

Yoon and Jo explore how direct and indirect access to corpora influences students' error correction and learning strategies in second language writing (Yoon, Jo, 2014). Their case study highlights the benefits of using corpora as a tool for improving writing accuracy and the development of self-regulated learning strategies. Similarly, Zahra and Abbas discuss the pedagogical implications of corpus-based approaches to English language teaching (ELT) in Pakistan, emphasizing the potential for corpora to enhance language learning by providing authentic language data (Zahra, Abbas, 2018). Both studies underline the value of corpus-based methods in promoting more effective language acquisition.

Incorporating corpus linguistics into the language classroom not only enhances learners' awareness of language nuances but also empowers them to develop more robust analytical skills. Such an evidence-based methodology fosters a learning environment where learners can interact with language data, thus becoming more autonomous and discerning users of English. Overall, the corpus-based approach serves as a foundational tool for promoting linguistic competence and communicative proficiency in English language learners.

### Conclusion

In conclusion, this study supports the effectiveness of corpus-based activities in improving English language proficiency, particularly in vocabulary and grammar. While students demonstrated a neutral to negative attitude toward corpus-based activities, they recognized their potential value in language learning.

These findings underscore the importance of integrating corpus-based activities into ELT curricula and suggest that more targeted interventions may be needed to improve student engagement and attitudes. With further research and refinement of teaching methods, corpus-based activities hold great promise for advancing English language education in Kazakhstan and beyond.

### References

- Alenizi, A., Adawi, R. (2024) Investigating the effectiveness of using corpus-based developed materials in vocabulary learning for Saudi EFL students. *Forum for Linguistic Studies*, 06, P. 721-740. <https://typeset.io/papers/investigating-the-effectiveness-of-using-corpus-based-4mg2609u6iol>
- Bernardini, S. (2002) Exploring New Directions for Discovery Learning. In B. Kettemann & G. Marko (Eds.), *Teaching and Learning by Doing Corpus Analysis*. Amsterdam: Rodopi, P. 165-182.
- Bernardini, S. (2004) Data-driven learning: Taking the computer out of the equation. *Language Learning & Technology*, 8(2), P. 165-173. <https://scholarspace.manoa.hawaii.edu>
- Boulton, A., Cobb, T. (2017) Corpus use in language learning: A meta-analysis. *Language Learning*, 67(2), P. 348-393. [https://www.researchgate.net/publication/313729217\\_Corpus\\_Use\\_in\\_Language\\_Learning\\_A\\_Meta-Analysis\\_Meta-Analysis\\_of\\_Corpus\\_Use\\_in\\_Language\\_Learning](https://www.researchgate.net/publication/313729217_Corpus_Use_in_Language_Learning_A_Meta-Analysis_Meta-Analysis_of_Corpus_Use_in_Language_Learning)
- Can, M. (2023) The impact of corpus-based approaches on the acquisition of academic vocabulary. *Language Learning*, 58(4), P. 358-373.
- Girgin, U. (2011) Corpus-based Activities at Lower Levels of EFL Proficiency: The Effectiveness of Using Concordance Lines on Grammar Learning.
- Hou, H. (2014) Teaching Specialized Vocabulary by Integrating a Corpus-Based Approach: Implications for ESP Course Design at the University Level. *English Language Teaching*, 7(5).
- Hunston, S. (2002) *Corpora in Applied Linguistics*. Cambridge: Cambridge University Press.
- Knoch, U. (2004). A New Look at Teaching Comparisons: A Corpus-based Study. *Journal of Language and Learning*, 2(2), P. 171-185.
- Leech, G., Candlin, C. (1997) *Teaching and Language Corpora*. Addison Wesley Longman.
- Liu, D., Jiang, P. (2009) Using a Corpus-Based Lexicogrammatical Approach to Grammar Instruction in EFL and ESL Contexts.
- O'Keeffe, A., McCarthy, M., Carter, R. (2007) *From corpus to classroom: Language use and language teaching*. Cambridge: Cambridge University Press. Cambridge Core
- Paker, T., Özcan, Y. (2017) The Effectiveness of Using Corpus-Based Materials in Vocabulary Teaching. *International Journal of Language Academy*. 5. P. 62-81. 10.18033/ijla.3494.
- Prodromou, L. (1996). Correspondence section. *ELT Journal*, 50(1), 88-89.



The history of Collins COBUILD (2020) <https://www.collinsdictionary.com/cobuild#corpus>

Tosun, C., Sofu, H. (2023) Challenges and benefits of corpus-based vocabulary instruction. *he Journal of Language Education*, Vol. 9, No. 3 (2023). P. 120-130.

Yoon, H., Jo, J. W. (2014) Direct and Indirect Access to Corpora: An Exploratory Case Study Comparing Students' Error Correction and Learning Strategy Use in L2 Writing. *Language Learning & Technology* 18(1), P. 96-117. [https://www.researchgate.net/publication/285986761\\_Direct\\_and\\_indirect\\_access\\_to\\_corpora\\_An\\_exploratory\\_case\\_study\\_comparing\\_students\\_error\\_correction\\_and\\_learning\\_strategy\\_use\\_in\\_l2\\_writing](https://www.researchgate.net/publication/285986761_Direct_and_indirect_access_to_corpora_An_exploratory_case_study_comparing_students_error_correction_and_learning_strategy_use_in_l2_writing)

Youssef, A. F. (2020) The Effectiveness of Corpus-Based Approach on Vocabulary Learning Gains and Retention in Saudi Tertiary EFL Context. *Journal of English Language Teaching and Applied Linguistics*, Volume 2, Issue 3, P. 1-15. The DOI is 10.32996/jeltal.2020.2.3.1.

Zahra, T., Abbas, A. (2018) Pedagogical Implications of Corpus-based Approaches to ELT in Pakistan. *Journal Of Education And Educational Development*, 5(2), 259 p.

*Information about the article / Мақала туралы ақпарат / Информация о статье.*

Entered the editorial office / Редакцияға түсті / Поступила в редакцию: 01.06.2024.

Accepted for publication / Жариялауға қабылданды / Принята к публикации: 23.12.2024.

© Yersultanova, G., Seidaliyeva, G., Kaiyrbekova, U., 2024

© A. Baitursynuly Institute of Linguistics, 2024